Vol. 55. Part 1 February 1964

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
BORIS SEMEONOFF

AIDED IN THE SELECTION OF PAPERS BY
C. BURT
AND

F. C. BARTLETT
E. A. BOTT
D. E. BROADBENT
R. S. CREED
J. DREVER

G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER

T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
M. D. VERNON

ASSISTANT EDITOR
R. M. McKENZIE

31

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W. 1
AMERICAN BRANCH, 32 EAST 57TH STREET, NEW YORK 22, N.Y.

Price Twenty-five Shillings net (U.S.A. $4.25)

Preparation of Papers for the British Journal of Psychology

The *British Journal of Psychology* subscribes to the *Fair Copying Declaration* of the Royal Society under the terms of which reproduction of material published in this Journal is permitted.

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. All papers, including reviews, must be typed (double spacing) on one side of the paper only and the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of most articles published from Vol. 47 (1956) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

2. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs etc. is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered caption of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following example:

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.
RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* 50, 83–96.

Note that World List abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

CUK-6303-31-249439

THE BRITISH JOURNAL OF PSYCHOLOGY

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

EDITED BY
BORIS SEMEONOFF

AIDED IN THE SELECTION OF PAPERS BY
C. BURT

AND

F. C. BARTLETT
E. A. BOTT
D. E. BROADBENT
R. S. CREED
J. DREVER
G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
M. D. VERNON

ASSISTANT EDITOR
R. M. McKENZIE

VOLUME 55, 1964

CAMBRIDGE
AT THE UNIVERSITY PRESS
1964

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York 22, N.Y.

Printed in Great Britain at the University Printing House, Cambridge
(Brooke Crutchley, University Printer)

CONTENTS OF VOLUME 55

PART 1. FEBRUARY, 1964

PAGE

Third-order factors in objective personality tests. By K. PAWLIK and R. B. CATTELL 1

The relative importance of proactive inhibition and degree of learning in retention of paired associate items. By P. B. WARR 19

The effects of the variation of two physical parameters on empirical difficulty in a perceptual maze test. By A. ELITHORN, D. JONES, MYFANWY KERR and D. LEE 31

Shape discrimination by rats: squares and rectangles. By N. S. SUTHERLAND and A. E. CARR 39

The effects of viewing conditions and instructions on shape judgements. By A. A. LANDAUER 49

Response instructions and the speed of relative judgements. 1. Some experiments on brightness discrimination. By R. J. AUDLEY and C. P. WALLIS 59

Acoustic confusions in immediate memory. By R. CONRAD. 75

Mimicry in mynas (*Gracula religiosa*): a test of Mowrer's theory. By B. M. FOSS 85

Consciousness and behaviourism. By J. D. KEEHN 89

Consciousness and behaviourism. A reply. By C. BURT 92

An electronic stimulus timer. By R. S. RODGER and R. K. QUIGG . . . 97

Publications recently received 103

Other publications received 119

PART 2. MAY 1964

Response instructions and the speed of relative judgements. II. Pitch discrimination. By C. P. WALLIS and R. J. AUDLEY 121

Constant errors in weight judgements as a function of the size of the differential threshold. By HELEN E. ROSS 133

Cues that elicit analytic-deductive methods in concept attainment. By WAYNE A. WICKELGREN 143

Semantic satiation of words and numbers By S. MESSER, L. A. JAKOBOVITS, R. KANUNGO and W. E. LAMBERT 155

The effects of context on the visual duration threshold for words. By JOHN MORTON 165

Imprinting and short-term retention. By W. SLUCKIN and K. F. TAYLOR . 181

PAGE

Further tests with the spiral after-effect of a theory of homoeostatic excitation–inhibition. By C. G. COSTELLO 189

Effects of sensory input and sensory dominance on severely disturbed, autistic children and on subnormal controls. By BEATE HERMELIN and N. O'CONNOR 201

Performance and level of activation in schizophrenics and normals. By P. H. VENABLES 207

The spiral maze. A psychomotor test with implications for the study of delinquency. By H. B. GIBSON 219

Publications recently received 227

PART 3. AUGUST 1964

OBITUARY NOTICE: ERNST KRETSCHMER 249

Personality variables in psychiatric classification: A symposium 253

- Specific factors and minor components. By P. SLATER 254
- Genetical factors in neurosis. By E. SLATER 265
- Personal continuity and psycho-pathological disruption. By G. A. FOULDS 269
- Personality in the diagnosis of hysteria. By J. G. INGHAM and J. O. ROBINSON 276
- Principles and methods of personality description, classification and diagnosis. By H. J. EYSENCK 284

Retroactive interference in the recall of simple sentences. By JACQUES MEHLER and GEORGE A. MILLER 295

Rate of presentation and order of recall in immediate memory. By MICHAEL I. POSNER 303

Changes in heart rate and performance as a result of loss of sleep. By D. W. J. CORCORAN 307

Perception of verticality in adult life. By ANN D. M. DAVIES and G. W. H. LEYTHAM 315

An analysis of the pursuit rotor learning of chronic psychotics. By P. L. BROADHURST and ANNE BROADHURST 321

A study of 200 school children by means of the test of family attitudes. By LYDIA JACKSON 333

An experimental study of imagery and its relation to abilities and interests. By K. RAY CHOWDHURY and P. E. VERNON 355

The titles of paperback books. By W. P. BROWN 365

Notes on communication in hospital. By G. W. H. LEYTHAM and R. B. HETHERINGTON 369

Publications recently received 371

PART 4. NOVEMBER 1964

PAGE

OBITUARY NOTICE: C. W. VALENTINE 385

Introjection: A form of social learning. By MICHAEL ARGYLE 391

Ignoring irrelevant information. By P. M. A. RABBITT 403

Sequential errors in a time-sharing task. By G. S. TUNE 415

Relation between perceptual and conceptual operations: Numerical distance and visual extent. By L. TEFT, S. WAPNER, H. WERNER and J. H. MCFARLAND 421

Information, acoustic confusion and memory span. By R. CONRAD and A. J. HULL 429

Transfer of training between space-oriented and body-oriented control situations. By M. HAMMERTON and A. H. TICKNER 433

Experimental studies in 'hypnotic' behaviour: suggested deafness evaluated by delayed auditory feedback. By T. X. BARBER and D. S. CALVERLEY . 439

Validation of the Maudsley personality inventory E scale. By T. M. CAINE and K. HOPE 447

An observed age difference in responses to a personality inventory. By F. I. M. CRAIK 453

Responses by the ferret to stimulus change. By R. N. HUGHES 463

A deuteranomalous artist. By R. W. PICKFORD 469

Season of birth, intelligence and personality measures. By ANN D. M. DAVIES 475

Baudouin on Jung. By CYRIL BURT 477

Publications recently received 485

Other publications received 507

INDEX OF AUTHORS

PAGE

Argyle,, M. 391
Audley, R. J. 59, 121
Barber, T. X. 439
Broadhurst, A. 321
Broadhurst, P. L. 321
Brown, W. P. 365
Burt, C. 92, 385, 477
Caine, T. M. 447
Calverley, D. S. 439
Carr, A. E. 39
Cattell, R B. 1
Chowdhury, K. R. 355
Conrad, R. 75, 429
Costello, C. G. 189
Craik, F. I. M. 453
Davies, A. D. M. 315, 475
Elithorn, A. 31
Eysenck, H. J. 249, 284
Foss, B. M. 85
Foulds, G. A. 269
Gibson, H. B. 219
Hammerton, M. 433
Hermelin, B. 201
Hetherington, R. R. 369
Hope, K. 447
Hughes, R. N. 463
Hull, A. S. 429
Ingham, J. G. 276
Jackson, L 333
Jakobovits, L. A. 155
Jones, D. 31
Kanungo, R. 155
Keehn, J. D. 89
Kerr, M. 31

PAGE

Lambert, W. E. 155
Landauer, A. A. 49
Lee, D. 31
Leytham, G. W. H. 315, 369
McFarland, J. H. 421
Mehler, J. 295
Messer, S. 155
Miller, G. A. 295
Morton, J. 165
O'Connor, N. 201
Pawlik, K. 1
Pear, T. H. 388
Pickford, R. W. 469
Posner, M. I. 303
Quigg, R. K. 97
Rabbitt, P. M. A. 403
Robinson, J. O. 276
Rodgers, R. S. 97
Ross, H. E. 133
Slater, E. 265
Slater, P. 254
Sluckin, W. 181
Sutherland, N. S. 39
Taylor, K. F. 181
Teft, L. 421
Tickner, A. H. 433
Tune, G. S. 415
Venables, P. H. 207
Vernon, P. E. 355
Wallis, C.P. 59, 121
Wapner, S. 421
Warr, P. B. 19
Werner, H. 421
Wickelgren, W. A. 143

INDEX OF PUBLICATIONS REVIEWED

Abercrombie, M. L. J. *Perceptual and Visuomotor Disorders in Cerebral Palsy* 498

Ajuriaguerra, J. de and others. *Problemes de Psycho-Linguistique* 373

Alcock, T. *The Rorschach in Practice* 233

Ammons, R. B. & Ammons, C. H. *The Quick Test (QT)* 117

Anastasi, A. *Fields of Applied Psychology* 486

Anstey, E. *Committees. How they Work and How to Work Them* 118

Argyle, M. *Psychology and Social Problems* 485

Astrup, C. *Schizophrenia: Conditional Reflex Studies* 110

Barker, C. E. *Psychology's Impact on the Christian Faith* 384

Barnett, S. A. *A Study in Behaviour* 106

Battin, R. R. & Haug, C. O. *Speech and Language Delay: A Home Training Program* 382

Bernard, E. E. & Kare, M. R. *Biological Prototypes and Synthetic Systems*, vol. I 113

Bernheim, H. *Hypnosis and Suggestion in Psychotherapy: the Nature and Uses of Hypnotism* 381

Binswanger, L. *Being-in-the World* 504

Bion, W. R. *Elements of Psycho-Analysis* 231

Buckner, D. N. & McGrath, J. J. (eds.). *Vigilance: a Symposium* 243

Bosch, G. *Der Frühkindliche Autismus* 499

Broadhurst, P. L. *The Science of Animal Behaviour* 245

British Journal of Industrial Relations, vol. I, nos. 1 and 2 118

Burnham, R. W., Hanes, R. M. & Bartleson, C. J. *Color: a Guide to Basic Facts and Concepts* 382

Cameron, N. *Personality Development and Psychopathology* 378

Chappell, V. C. (ed.) and others. *The Philosophy of Mind* 105

Cohen, J. (ed.). *Readings in Psychology* 487

Cooley, W. W. & Lohnes, P. R. *Multivate Procedures for the Behavioral Sciences* 115

Cooper, J. B. & McGaugh, J. L. *Integrating Principles of Social Psychology* 107

Cornsweet, T. N. *The Design of Electric Circuits in the Behavioural Sciences* 383

Critchley, M. *Developmental Dyslexia* 282

David, H. A. *The Method of Paired Comparisons* 115

Davson, H. (ed.). *The Eye*. Vol. II: *The Visual Process* 246

Davson, H. (ed.). *The Eye*. Vol. IV: *Visual Optics and the Optical Space Sense* 246

Donaldson, M. *A Study of Children's Thinking* 241

Eickstedt, E. F. von. *Ursprung Und Entfaltung Der Seele* 505

Eissler, K. R. *Leonardo da Vinci* 380

Etkin, W. *Social Behavior and Organization Among Vertebrates* 497

Eysenck, H. J. (ed.). *Experiments in Motivation* 496

Festinger, L. *A Theory of Cognitive Dissonance* 228

Festinger, L., Schachter, S. & Back, K. *Social Pressures in Informal Groups* 237

Fine, R. *Freud: a Critical Re-evaluation of his Theories* 501

Flavell, J. H. *The Developmental Psychology of Jean Piaget* 107

Ford, D. H. & Urban, H. B. *Systems of Psychotherapy* 500

Fordham, M. *Contact with Jung* 233

Foss, B. M. (ed.). *Determinants of Infant Behaviour: II* 239

Fraisse, P. *The Psychology of Time* 104

Furlong, E. J. *Imagination* 375

Gabriel, J. *Children Growing Up* 494

Goffman, E. *Behaviour in Public Places* 238

Gordon, J. E. *Personality and Behavior* 236

Greidanus, J. H. *Fundamental Physical Theory and the Concept of Consciousness* 227

Grünhut, M. *Probation and Mental Treatment* 248

Häfner, H. *Psychopathen: Daseinsanalytische Untersuchungen Zur Struktur Und Verlaufsgestalt Von Psychopathien* 499

Hahn, M. E. *Psychoevaluation* 118

Handy, R. & Kurtz, P. *A Current Appraisal of the Behavioural Sciences* 506

Hans, N. *The Russian Tradition in Education* 103

Harms, E. & Schreiber, P. (eds.) *Handbook of Counselling Techniques* 503

Hécaen, H. & Angelergues, R. *La Cécité Psychique* 111

Henderson, Sir David K. *The Evolution of Psychiatry in Scotland* 380

Hernandez Peon, R. (ed.). *The Physiological Basis of Mental Activity* 492

Hollander, E. P. & Hunt, R. G. (ed.). *Current Perspectives in Social Psychology* 376

Homans, G. C. *Social Behavior—Its Elementary Forms* 236

Hsu, F. L. K. *Clan, Caste and Club* 490

Humphrey, G. (ed.). *Psychology Through Experiment* 105

Humphrey, G. & Coxon, R. V. *The Chemistry of Thinking* 247

Jalota, S. *Experiments in Psychology* 488

Jaspers, K. *General Psychopathology* 112

Jung, C. G. *Memories, Dreams, Reflections* 233

Kadinsky, D. *Strukturelemente Der Personlichkeit* 234

Kantor, J. R. *The Scientific Evolution of Psychology. I* 374

Keehn, J. D. *The Prediction and Control of Behavior* 106

Kline, M. V. *Clinical Correlations of Experimental Hypnosis* 381
Kline, M. V. (ed.). *The Nature of Hypnosis* 113
Knapp, P. H. (ed.). *Expression of the Emotions in Man* 491
Kroger, W. S. *Clinical and Experimental Hypnosis* 113
Laing, R. D. & Esterson, A. *Sanity, Madness and the Family*, vol. I 504
Lawrence, D. H. & Festinger, L. *Deterrents and Reinforcement* 228
Leuner, H. *Die Experimentelle Psychose* 499
Lewis, M. M. *Language, Thought and Personality in Infancy and Childhood* 240
Lidz, T. *The Family and Human Adaptation* 379
Luce, R. D., Bush, R. R. & Galanter, R. (eds.). *Handbook of Mathematical Psychology*, vol. I, chapters 1–8 245
Luce, R. D., Bush, R. R. & Galanter, E. (eds.). *Readings in Mathematical Psychology*, vol. I 245
Mackenzie, M. *Psychological Depression* 247
Malan, D. H. *A Study of Brief Psychotherapy* 110
Malmberg, B. *Structural Linguistics and Human Communication* 371
Mays, J. B. *Crime and the Social Structure* 248
Meng, H. & Freud, E. L. (eds.). *Psycho-Analysis and Faith: the Letters of Sigmund Freud and Oskar Pfister* 501
Messick, S. & Ross, J. (eds.). *Measurement in Personality and Cognition* 235
Michael, J. *Laboratory Studies in Operant Behavior* 383
Miller, G. A. *Psychology: The Science of Mental Life* 485
Mowbray, R. M. & Rodger, T. F. *Psychology in Relation to Medicine* 109
Newson, J. & Newson, E. *Infant Care in an Urban Community* 108
O'Connor, N. & Hermelin, B. *Speech and Thought in Severe Subnormality* 242
O'Neil, W. M. *An Introduction to Method in Psychology* (second edition) 106
Pearson, E. S. *Tables for Testing Significance in a* 2 × 2 *Contingency Table* 115
Pickford, R. W. *Projective Pictures* 117
Piotrowski, Z. A. & Rock, M. R. *The Perceptanalytic Executive Scale* 495
Pitcher, E. G. & Prelinger, E. *Children Tell Stories* 495
Rachman, S. (ed.). *Critical Essays on Psychoanalysis* 502
Rheingold, H. L. (ed.). *Maternal Behavior in Mammals* 498
Roback, A. A. *Aspects of Applied Psychology and Crime* 489
Rubenowitz, S. *Emotional Flexibility-Rigidity as a Comprehensive Dimension of Mind* 235
Sainsbury, P. & Kreitman, N. (eds.). *Methods of Psychiatric Research* 115
Sarnoff, I. *Personality: Dynamics and Development* 231
Sauerbruch, F. & Wenke, H. *Pain. Its Meaning and Significance* 246
Scheerer, C. (ed.). *Cognition: Theory, Research, Promise* 491
Schneck, J. M. *Hypnosis in Modern Medicine* (third edition) 113
Schon, D. A. *Displacement of Concepts* 229
Scott, W. A. & Wertheimer, M. *Introduction to Psychological Research* 115
Simon, B. & Simon, J. (eds.). *Educational Psychology in the U.S.S.R.* 103
Smart, J. J. C. *Philosophy and Scientific Realism* 373
Smelser, N. J. & Smelser, W. T. *Personality and Social Systems* 377
Sokolov, Ye. N. *Perception and the Conditioned Reflex* 493
Solnit, A. J. & Provence, S. A. (eds.). *Modern Perspectives in Child Development* 379
Staats, A. W. & Staats, C. K. *Complex Human Behavior* 494
Stafford-Clark, D. *Psychiatry for Students* 381
Stevenson, H. W. (ed.). *Child Psychology. The 62nd Yearbook of the National Society for the Study of Education. Part I* 240
Stuart, A. *Basic Ideas of Scientific Sampling* 115
Summerfield, A. (ed.). *British Medical Bulletin: Experimental Psychology* 243
Super, D. E. & Crites, J. O. *Appraising Vocational Fitness by Means of Psychological Tests* 116
Tarachow, S. *An Introduction to Psychotherapy* 110
Thorpe, W. H., F.R.S. *Learning and Instinct in Animals* (second edition) 106
Thouless, R. H. *Experimental Psychical Research* 244
Tomkins, S. S. *Affect Imagery Consciousness. Vol.* I 230
Tomkins, S. S. *Affect Imagery Consciousness. Vol.* II 376
Tomkins, S. S. & Messick, S. (eds.). *Computer Simulation of Personality* 114
Vasiliev, L. L. *Experiments in Mental Suggestion* 244
Vygotsky, L. S. *Thought and Language* 103
Wepman, J. M. & Heine, R. W. (eds.). *Concepts of Personality* 496
Werner, H. & Kaplan, B. *Symbol Formation* 492
Westby, G. *Behaviour Theories and the Status of Psychology* 384
Whiting, B. B. (ed.). *Six Cultures: Studies of Child Rearing* 108
Wilner, D. M. and others. *The Housing Environment and Family Life* 109
Wooton, B. *Crime and the Criminal Law* 248

Brit. J. Psychol. (1964), 55, 1, *pp.* 1–18
Printed in Great Britain

THIRD-ORDER FACTORS IN OBJECTIVE PERSONALITY TESTS*

By KURT PAWLIK
University of Vienna

AND RAYMOND B. CATTELL
University of Illinois

Intercorrelations of second-order factors in objective personality tests were pooled from three researches and subjected to a new analysis in terms of third-order personality factors. Application of Guttman's criterion allowed for the extraction of three factors which, after combined analytical and visual rotation for oblique simple structure, could be identified as 'Immature Self-Centred Temperament (*v.* Mature Environment-Oriented Stolidity)', 'Restrained Acceptance (*v.* Unrestrained Neglect) of External Norms', and 'High (*v.* Low) Self-Assertion', respectively. A brief review of first- and second-order factor structure in objective personality tests is given in the introductory section.

This paper reports recent findings on personality structure, at the level of third-order factors, in objective personality tests. In doing so it brings out certain principles of both methodological and theoretical interest.

As is generally recognized, the fruitfulness of further analysing oblique primary factors into higher-order constructs is directly related to the degree of accuracy and precision achieved in locating the former—just as the third storey of a house has only the stability supplied by the first. Exploration of the higher-order factor structure has therefore been deliberately postponed until the simple structure of the primary factors should become thoroughly replicated. The evidence on some twenty such first-order factors in objective personality tests is summarized in a monograph by Hundleby, Pawlik & Cattell (1964), while independent second-order studies have meanwhile checked well, as shown by Cattell, Knapp & Scheier (1961). Before presenting the findings of the present research it is desirable, however, to comment briefly on the methodology of those studies and to give the reader a brief sketch of the nature of the primary and second-order factors upon which this analysis is based.

I. PERSONALITY RESEARCH: ESSENTIAL CONTRASTS OF UNIVARIATE AND MULTIVARIATE EXPERIMENT

It has been voiced for some years that neither clinical psychology, on the one hand, nor the classical univariate experiment on the other have sufficed to bring personality theory to a satisfactory development. The former proceeds holistically enough to encompass the intricate laws which probably exist, but it is typically pre-metric in observation and thus meets unsurmountable problems if quantitative laws are sought. The latter labours under the systematic disability that concepts which involve patterns—and most personality concepts do—can never be operationally represented by simple measures to yield clear experimental answers.

---

* This investigation was supported in part by U.S. Public Health Service research grant M-1733, National Institute of Mental Health.

By contrast, the methodological 'axioms' of the multivariate experimentalist, employing factor analysis, canonical correlations, etc., can be stated as follows:

(1) The domain of individual behaviour differences is better viewed as *multivariate* in nature, since the number of distinct, reliably measurable behaviour differences is very large (however presumably finite) and no single variable or group of variables sampled from a sub-domain of this universe has any claim exhaustively to represent the total 'personality sphere'. Progress is unlikely to be made without reduction to major dimensions which all experimenters can use in common.

(2) Individual behaviour differences are not only large in number but their complex modes of mutual interdependence demand a model of behaviour implying multivariate causation of each observable variable, i.e. any kind of one-to-one correspondence between observable variables and causal entities is least likely. Or in other words: each behaviour phenomenon is, most likely, determined by several causes (principle of multiple causation of behaviour), as each cause is most likely to affect more than one behavioural phenomenon (principle of multiple effect of behaviour causes).

These two 'axioms' for personality research have really only been rediscovered by modern multivariate personality research—their appropriateness was realized long before factor analysis was formulated. They seemingly apply to human behaviour in a very general sense, and the researcher can easily violate them in any of the three following respects (which loosely correspond to the three principal phases of any scientific research): data collection (choice of an appropriate experimental design), data analysis (choice of a suitable model of analysis), and data interpretation (development of theoretical models in accordance with these axioms).

Historically—and thus traditionally—psychological research has been less cognizant of the multivariate nature of human behaviour than the psychological 'practitioner' might often have desired. Following perhaps all too closely the parental footsteps, psychology as the youngest child of the physical sciences (and in particular, general experimental psychology) has been striving after a theory of behaviour whose structure was still deeply indebted to theoretical models deemed successful in traditional physiology or macrophysics, and whose validity seemed best explorable by *univariate* experimentation. By this we mean an experimental design which explores *one single dependent variable* Y_i at a time, studying it as a function of one or more variables $X_1, X_2, \ldots, X_n$, which are regarded as independent variables.

Thus in univariate experimentation one facet or even only one single cell of the hypothetical data matrix is explored at a time; and regardless of how often this experiment is repeated, with different dependent variables $Y_1, Y_2, \ldots, Y_m$ each time, the experimenter will still be lacking information on the interrelationship between any behavioural data Y_i and Y_j or on the joint determination of Y_i and Y_j by the causal entity X.* Not only does this design exclude the collection of otherwise important information but it eventually also streamlines psychological theory to fit the heuristic model implied in univariate experimental designs.

---

* Cf. Mood (1950, p. 358): 'Most experimental work today is based on the rule: "Keep all variables constant but one", an ancient and erroneous dictum which guarantees a high degree of inefficiency. One well-designed experiment, taking account of all relevant factors, is worth dozens or even hundreds of experiments which study one factor at a time keeping the others constant.'

If, in this classical type of psychological experiment, a significant variation is observed in Y_i upon a produced variation in a given X_j this is typically interpreted as an effect upon Y_i caused by X_j. By its very nature, this univariate (and at the same time: unidirectional) experimental design is likely to result in a univariate—and unidirectional—theory of behaviour. If one so wishes one may regard traditional stimulus-response theory of behaviour as both a constant reinforcer and a necessary outgrowth of this experimental method. The 'artificial' situation in which the effect of influences is studied one at a time, however, causes the investigator to overlook inherent relations between two or more variables; in addition, the different 'factors' $X_1 \ldots X_n$ are set orthogonal to each other in the design of the experiment and treated as orthogonal in the subsequent analysis. In other words: even in the most complex univariate designs the distinct 'factors' or 'causes' X_i are made independent or uncorrelated; only their effects are allowed to correlate ('interaction' in analysis of variance terminology). As a consequence, inferences drawn from data obtained with independent factors X_j will be systematically biased if, in 'nature', the behaviour causes are actually correlated. Such a bias will be even larger if the chosen (and experimentally produced) variation in X_j constitutes an unrepresentative sample from the natural variation in X_j (cf. Brunswik's (1956) notion of 'systematic' *v.* 'representative' experimental designs).

The foregoing evaluation of univariate experimental research techniques has been made explicitly from the standpoint of personality research and no further. It leads to the following summary:

(1) By restricting the experiment to one single behavioural variable Y_i, information about any cross-relationship with other behavioural variables Y_j is not only unknown but also constantly neglected in interpreting the results. If Y_i has been chosen to represent a certain hypothetical construct (e.g. anxiety) conclusions drawn from the experiment are very likely to be in error because a single behavioural variable will measure this construct typically with only low construct validity (Cattell & Scheier, 1958, 1961).

(2) Univariate experiments introduce (an assumed rather than a discovered) unidirectionality into the investigated cause–effect relationships pertaining to a behavioural datum Y_i and a set of 'independent' variables X_j.

(3) In *deciding* about the nature of the variables X_j to be included in a study as well as their variation to be produced in the experiment, rather than *assessing* them under natural variation, the researcher may create an unrepresentative image of nature—this even more so if, in the course of the experiment, the interrelationship between any two 'factors' X_i and X_j is not modelled after their natural correlation but set orthogonal on *a priori* grounds.

By contrast, in *multivariate experimental designs* it is not a single behaviour variable Y_i that is studied, but a set of such variables $Y_i \ldots Y_m$; the latter constitutes a sample of variables, either selected to sample a complete domain of behavioural phenomena or chosen on the basis of previous results for the purpose of multivariate measurement of one or more hypothetical constructs (factors). Unlike univariate experimentation, multivariate research does not decide at the outset which variables are 'dependent' (effected) and 'independent' (effectors) and which hypothetical constructs are most likely to prove useful in the course of explaining overt behaviour.

The answers to these questions are aims rather than premises of a multivariate experiment. Stating it more explicitly, the aims of multivariate personality research are in essence threefold:

(1) The construction of measurement instruments such that the domain of behaviour differences, the 'personality sphere', can be sampled representatively by way of a limited set of variables.

(2) The discovery of those 'factors' or 'causal entities' that underly the observable behaviour variation, thus delineating a model of personality which expresses observable behaviour as a function of these factors.

(3) The further investigation of these factors, in order to arrive at a substantive theory over and above the variable-factor relationships.

These aims are perpetuated by the multivariate researcher in what may be described as a four-stage procedure as outlined schematically in Table 1.

Table 1. *The four principal stages in multivariate personality research*

	Methods
1. Delineation of behaviour universe ('personality sphere') in terms of observable traits	Listing of 'surface traits' (Cattell, 1946), relying upon: free behaviour observation in real-life situations, logical condensation from all existing verbal expressions descriptive of personality characteristics (e.g. Allport & Odbert, 1936)
2. Construction of a sample of surface trait measures to represent this universe	Development of suitable assessment instruments, e.g. objective behaviour tests, from which these traits can be measured. In addition to pure behavioural measures, non-behavioural (physiological, physical, environmental) variables should be considered as well
3. Discovery of the primary factors underlying the behavioural data	Multiple factor analysis of the obtained variable sample through *several* replicative studies (using replicability of a factor as criterion of its behavioural generality)
4. Exploration of meaning and behavioural significance of the obtained factors	Testing hypotheses about the discovered factors or 'source traits' (Cattell, 1946), accumulating information on their mechanisms of development (e.g. nature/nurture results, age development) and operation

The following considerations are essential to this multivariate approach to the study of personality: (*a*) Research is not centred, from the very beginning, on preselected classes of 'dependent', and 'independent' variables—the factors or causes determining overt behaviour are obtained from analysing a representative sample of the total data matrix into more basic entities underlying the observed stimulus–response contingencies. (*b*) Instead of experimentally introducing a (potentially unrepresentative) variation in the factors of possible influence upon behaviour, natural inter-individual variation in behaviour, under fixed stimulus conditions (or intra-individual, as in *P*-technique (Cattell, 1946, p. 96ff.)), constitutes the basic data matrix. (*c*) Behavioural phenomena are studied in their mutual interrelationships rather than singly.

According to the 'strategy' of research laid out in Table 1, factor analysis is only part of a multivariate research programme in personality. Once factors have been located and rendered accurately measurable by objective test batteries or questionnaire batteries, there is no need for factor analytic skills in their use in psychological experiments. Thus step 4 in Table 1 is primarily non-factor analytic, and factor analysis only enters into it in so far as this step also includes the analysis of the primaries in terms of higher-order factors. This latter undertaking, of which the present study is an example, is likely to yield the double gain of discovering certain broader interactions not yet known and of helping in the interpretation of the primaries themselves.

II. THE PRESENT RESULTS ON REPLICATED FIRST-ORDER AND SECOND-ORDER FACTORS IN OBJECTIVE PERSONALITY TESTS

In a systematic multivariate approach to studying personality, some 20 years ago, Cattell (1945) started on the first exploratory work on the delineation of what he called the 'personality sphere', the universe of all conceivable personality characteristics. After an initial condensation of the 'surface trait' list thus obtained, subsequent research—corresponding to stage 2 in Table 1—was directed towards the development of a core set of assessment instruments for measuring each one of the personality surface traits contained in this list. In fact, this variable selection and construction work was done concurrently in all three distinguishable media of behaviour assessment: real-life data *in situ* (L) (including records and behaviour ratings (BR)), questionnaire responses (Q), and objective tests (OT). Factorization of the included surface traits into a smaller set of personality factors (or 'source traits') was accomplished in all three media, with good cross-media convergence of the factors obtained (Cattell & Saunders, 1950, 1955; Cattell & Beloff, 1956).

During the last 15 years, research by Cattell and his co-workers has been devoted largely to the two most promising of the three media, to questionnaire responses and objective test data. A first extensive summary of the respective research findings and general principles was given in 1957 (Cattell, 1957*a*), a more completely documented, up-to-date report on the factor structure in objective personality tests is presented in a recent monograph (Hundleby *et al.* 1964). A description of the more than 300 objective tests used in the OT factorizations is given by Cattell & Warburton (1964). The briefest possible outline of these factors will be given here to familiarize the reader with the principal findings on these first-order and second-order factors in objective personality tests, since the psychological meaning of the third-order factors can only be evaluated from a proper understanding of the factor structure obtained at the preceding levels of analysis.

(*a*) *The nature of the primary personality factors in objective tests*

In their recent review of the evidence on objective personality test dimensions, Hundleby *et al.* (1964) survey the results of more than 30 independent multivariate researches in which the factor structure of objective personality tests has been investigated. Studies included in this survey are limited to those which have enough identical test variables in common so that cross-studies factor matching could be objectively evaluated. Thus the final list of included researches contains, in addition to studies by Cattell and his co-workers, factor analytic researches by Ryans (1938), Thurstone (1938, 1944), Thornton (1939), Brogden (1940), Eysenck (1947, 1952*a*, *b*), and a few others. No detailed reference will be made here to these analyses, and the reader is referred to the Hundleby *et al.* (1964) monograph for more specific information on the nature, design, and results of these studies.

Re-examination of the matching of factors between the different analyses revealed a series of twenty-one primary objective personality test factors which showed adequate cross-studies consistency and of which all but one have been repeatedly identified in both adult and child populations. Following Cattell's early suggestion (Cattell, 1957*b*), these twenty-one factors have been indexed in a Universal Index (U.I.) series, starting at U.I. 16 (U.I. 1 to U.I. 15 correspond to the principal factors in ability tests as reviewed by French, 1951). For each of these twenty-one factors the loading pattern (in terms of 'salient' test variables) has been integrated from all the researches in which it could be identified. The resulting average loading pattern, which comprises up to sixty and more consistently loading test variables for some of the factors, defines each factor in terms of observable test behaviour and constitutes the basis for interpreting the factors.

In Table 2 these factors are listed by their U.I. number and current interpretative title. The degree of factor analytic replication is indicated both by the number of studies in which a factor has emerged (for children and adult subjects separately) and by the total number of subjects on which these studies have been based. In the right-hand column of Table 2 a brief psychological description is given for each factor, always referring to the positive factor pole (all factors are bipolar). In evaluating these descriptions the reader should keep in mind that they represent a first psychological abstraction from the averaged loading pattern, each descriptive phrase being based upon several, interpretatively convergent, test variables.

Some of these personality source traits have been described previously by other researchers under different titles. U.I. 23, for example, corresponds to Eysenck's neuroticism factor (in reversed direction), U.I. 25 closely resembles Eysenck's psychoticism factor (again in reversed direction), and factor 2 in Eysenck's cyclothymia-schizothymia analysis (Eysenck, 1952*a*) is in good agreement with the pattern of factor U.I. 21. U.I. 17 matches with Thurstone's (1944) factor $F-$ (Slowness of Perception), U.I. 19 with Thurstone's factor $E+$ (Suppressive Manipulation of Configurations), and U.I. 22 with Thurstone's $C+$ (Speed of Reaction). As to the last three factors, their pattern is no longer that of plain ability factors as they appeared in Thurstone's work, for the broader sampling of variables from a wider domain has now revealed them as temperament factors with the ability component being only a part of the total pattern.

Additional, non-factor analytic, results on these factors are presented schematically in Table 3. The column headings and entries should be understood as follows:

(1) *Evidence on state factors*

A factor obtained from analysing intra-individual response variation over time (such as in *P*-technique) instead of inter-individual behaviour differences is more likely to be a state factor rather than a trait. Replicative *P*-technique analyses (see Cattell, 1955*a*, 1957*a*, p. 633ff.) and a study using the 'incremental *R*-technique' design (Cattell & Scheier, 1960) did in fact show ten primary state factors in objective personality tests. An asterisk in the first column of Table 3 indicates that this factor

Table 2. *The twenty-one principal personality factors in objective tests*

U.I. no. of factor	Factor title	Factor analytic replication rate						Psychological description of the positive factor pole
		Adults		Children		Total		
		No. of studies	No. of subjects	No. of studies	No. of subjects	No. of studies	No. of subjects	
U.I. 16	Assertiveness *v.* Submissiveness	10	2212	5	725	15	2937	Fast natural tempo and fast speed of motor and cognitive performance; carefulness; high self-assertion and frustration tolerance; upper class tastes
U.I. 17	Timidity *v.* Adventurousness	13	2589	3	428	16	3017	High cautiousness and timidity; readiness to inhibit overt behaviour under threat; less alert
U.I. 18	Smartness *v.* Thoroughness	9	2070	4	525	13	2595	More rapid, but superficial judgements; faster speed on motor tasks; personally less secure; more susceptible to social norms
U.I. 19	Critical Independence *v.* Subduedness	13	2655	5	754	18	3409	Controlled, competent exactness; low distractability; self-complacent criticalness of others; flexible thinking, low social suggestibility
U.I. 20	Comention (Acceptance of Culture) *v.* Unresponsiveness	9	2023	5	657	14	2680	High acquiescence and tendency to conform with group norms; high sensitivity to punishment and disapproval
U.I. 21	Exuberance *v.* Quietness	11	2341	5	754	16	3095	Imperviousness to social suggestions; faster speed of social judgements, and faster perceptual speed; high verbal and ideational fluency
U.I. 22	Cortertia (Cortical Alertness) *v.* Pathemia	12	2437	5	754	17	3191	High immediate responsiveness, but low endurance; fast speed of reaction; tense restlessness and high level of activation; fast perceptual speed
U.I. 23	Ego Strength *v.* Neurotic Debility	10	1611	4	625	14	2236	Lack of rigidity; high ability to concentrate; high endurance of difficulty

Table 2. (*cont.*)

U.I. no. of factor	Factor title	Factor analytic replication rate						Psychological description of the positive factor pole
		Adults		Children		Total		
		No. of studies	No. of subjects	No. of studies	No. of subjects	No. of studies	No. of subjects	
U.I. 24	High Anxiety *v.* Low Anxiety	8	1383	6	854	14	2237	High manifest anxiety; increased susceptibility to annoyance and embarrassment; insecurity (feelings of guilt and self-depreciation); high drive tension level
U.I. 25	Realism *v.* Psychotic Tendency	8	1254	5	754	13	2008	Low cognitive and perceptual imaginative power; high accuracy and speed on well-structured (conventional) tasks; reality-directed attitudes
U.I. 26	Narcissism *v.* Non-Involvement	8	1376	5	725	13	2101	Pervasive efficiency and sophistication of response; high persistence; low self-evaluation, but high self-concern; high sensitivity to threat; good immediate control
U.I. 27	Grudgingness *v.* Trustingness	4	454	1	229	5	683	Low self-esteem; low aspiration level and test motivation
U.I. 28	Rigidity of Superego *v.* Casualness	8	1427	4	625	12	2052	Low psycho-physical 'momentum'; lack of self-confidence and of steadfastness of opinion; increased negativistic criticalness
U.I. 29	Wholeheartedness *v.* Unco-operativeness	7	1816	3	393	10	2209	High speed and co-operativeness of response; little effect resulting from discomfort and threat; over-reacting
U.I. 30	Stolidness *v.* Tenseness	5	1104	3	461	8	1565	Stolid temperament (slow and less careful); little fluctuation of attitudes; restricted emotionality (aloof, 'dry')
U.I. 31	Wariness *v.* Slaphappiness	4	950	2	361	6	1311	High stability of attitudes; high cautiousness; determined self-control; slow natural tempo
U.I. 32	Extraversion *v.* Introversion	5	1090	4	574	9	1664	'Extraversion'; optimistic self-confidence; high perceptual and cognitive fluency; social interests

Table 2. (*cont.*)

U.I. no. of factor	Factor title	Factor analytic replication rate: Adults: No. of studies	Adults: No. of subjects	Children: No. of studies	Children: No. of subjects	Total: No. of studies	Total: No. of subjects	Psychological description of the positive factor pole
U.I. 33	Pessimism (Dourness) *v.* Optimism	3	368	1	164	4	532	Pervasive pessimism; depressed lack of self-confidence; conservative and anxious approach to environment
U.I. 34	Autism *v.* Conventionality	4	381	2	209	6	590	Reasoning more 'emotional' than 'rational'; indifferent regarding social norms; high imaginative fluency; interest in the unusual rather than the usual
U.I. 35	Laziness *v.* Enthusiasm	3	340	1	80	4	420	Low level of cognitive performance; sluggishness and torpidity of behaviour, lacking incisiveness; lack of fine motor control
U.I. 36	Strong Self-Sentiment *v.* Lack of Foresightedness	3	309	—	—	3	309	High self-sentiment; foresighted restrictiveness

resembles closely one of the known state factors of personality; the implication is that the factor in question at least partly represents state rather than stable trait variance.

(2) *Criterion relationships*

The entries represent the signs of statistically significant criterion associations; for the clinical criteria a positive sign indicates a significantly higher mean score in the first-mentioned criterion group. The correlations with pilot training criteria (Cattell *et al.* 1953; Knapp & Most, 1960) and with the normal-neurotic dichotomy (study R 9-Ta in Cattell & Scheier, 1961, p. 66; Cattell, 1961) are each compiled from two studies, the remaining criterion validities are each based on one analysis (comparisons with a delinquent sample, Cattell, 1961; college achievement, Pawlik, 1961).

(3) *Nature–nurture evidence*

The results reported in this column have been obtained in pre-adolescent children (Cattell, Stice & Kristy, 1957). An 'E' indicates primarily environmental, and 'H' primarily hereditary determination of a factor.

(4) *Age trend results*

The entries in this column indicate the increase (positive sign) or decrease (negative sign) in mean factor score over the age period indicated; each value is expressed in units of the within-age factor score standard deviation (Cattell, 1957*a*, pp. 612ff.)

Blank cells in Table 3 indicate non-significant relationships; the ? denotes conflicting results between independent analyses. A crossed-out cell (×) shows that no results are available as yet on this factor.

Table 3. *Additional, non-factor analytic evidence on primary factors in objective personality tests*

		Criterion relationships						
	Identifiable also as state factor	Success in pilot training	Normals *v.* Neurotics	Non-delinquents *v.* Delinquents	Delinquents *v.* Neurotics	College achievement	Nature-nurture evidence	Age trend: 9–15 years
U.I. 16		+	+		+		E	$+1/8\sigma$
U.I. 17	*	−					E	$+1\sigma$
U.I. 18		+					×	×
U.I. 19	*		+			+	E	$+3/8\sigma$
U.I. 20		+				+	H	$+1/4\sigma$
U.I. 21		+	+	×	×	×	H	$-1/2\sigma$
U.I. 22	*		+	×	×	×	E	$+3\frac{1}{2}\sigma$
U.I. 23	*	+	+				E	$+7/8\sigma$
U.I. 24	*		−	−		−	×	×
U.I. 25		?	+	+	−		×	×
U.I. 26	*		+	+			E	$+7/8\sigma$
U.I. 27		×				−	×	×
U.I. 28	*		?	+			E	$-3/8\sigma$
U.I. 29			+	×	×	×	E	$-3/8\sigma$
U.I. 30				+			×	×
U.I. 31				×	×	×	×	×
U.I. 32			+	+			×	×
U.I. 33		×		+		−	×	×
U.I. 34		×				−	×	×
U.I. 35	*	×		−		−	×	×
U.I. 36		×	×	×	×	×	×	×

(*b*) *The recent evidence on second-order factors in objective personality tests*

Cattell *et al.* (1961) recently reported results on second-order personality factors in objective tests which, like those in the above discussion of the primary factors, were based on several replicative analyses. Their summary is in fact based upon a total of five independent studies (comprising in all 1237 adult subjects). Four of these researches were subsequent second-order factorizations of previous studies, already included in the averaged loading pattern of the primaries, in which the second-order analysis started from the primary factor intercorrelation matrix computed from the exact reference vector intercorrelations. In one study (Knapp's sample of 315 U.S. Navy Submarine School candidates) the intercorrelations of the primaries have been obtained from direct measurement of the first-order factors.

At present seven second-order factors are well-replicated across different analyses. They are listed in Table 4 by their index number and descriptive title. The primaries loading each of these factors are also reported,* the principal salients being italicized. The sign after each U.I. number indicates the sign of the loading of this primary factor on the given second-order factor.

---

* U.I. 1+ refers, as in French's index system (French, 1951), to a factor of general intellectual ability.

In terms of first-order factor variance accounted for, factor F(T)I, called *Tied Socialization v. Absence of Cultural Introjection*, is the largest of the seven second-order factors. It is interpreted as a personality source trait which determines the degree to which a person will internalize the rules, demands, and norms existing in his culture. The finding of Cattell & Scheier (1961) that this factor correlates negatively with the dichotomy normal–neurotic further substantiates this inference from the primary factors loading on F(T)I.

Table 4. *The seven known second-order factors in objective personality tests*

Factor index	Factor title	First-order salients
F(T)I	Tied Socialization *v.* Absence of Cultural Introjection	*U.I.20*+, *U.I.28*+, *U.I.1*−, *U.I.19*−, U.I. 25+, U.I. 35+, U.I. 32+
F(T)II	Expansive Ego *v.* History of Difficulty in Problem Solving	*U.I.16*+, *U.I.1*+, *U.I.19*+, U.I. 23−, U.I. 36+, U.I. 18−
F(T)III	Temperamental Ardour *v.* Temperamental Apathy	*U.I.21*+, *U.I.20*+, *U.I.1*−, *U.I.19*+, U.I. 27−, U.I. 32+
F(T)IV	High Educated Self-Consciousness *v.* Low Self-Consciousness	*U.I.22*+, *U.I.18*+, *U.I.25*−, U.I. 36+, U.I. 30−, U.I. 29−, U.I. 33+
F(T)V	History of Inhibiting, Restraining Environment *v.* Laxness	*U.I.17*+, U.I. 23+, U.I. 31+
F(T)VI	Narcissistic Development *v.* Environmental Contact and Investment	*U.I.26*+, *U.I.27*+, U.I. 34+
F(T)VII	High Tension to Achieve and Controlled Drive Tension Level *v.* Low Tension to Achieve	*U.I.24*+, *U.I.18*+, U.I. 30−

F(T)II, interpreted as *Expansive Ego v. History of Difficulty in Problem Solving*, shows several aspects of what the clinician might refer to as 'ego strength'. The self-determination, ego expansiveness, and high will strength in factor F(T)II is currently explained by a developmental hypothesis, in the sense that these attributes of ego strength result from a record of successful drive satisfaction in an individual's history.

The three principal first-order salients of F(T)III, U.I. 1, U.I. 20, and U.I. 21, are all largely constitutional (hereditarily determined) factors. Factor F(T)III, *Temperamental Ardour v. Temperamental Apathy*, therefore appears as a second-order source trait primarily determined by heredity. Its major psychological characteristic is a voluntary exuberance of response together with a tendency towards free (unrestrained) emotional expression. The loading of first-order factor U.I. 20 is somewhat puzzling in this connexion; however, U.I. 20 loaded highly on F(T)III in only two of the five second-order analyses in adults, and recent work by the present authors on the second-order factor structure in pre-adolescent children did not substantiate U.I. 20 as a salient of F(T)III. The qualities of low personal restraint and uninhibited temperamental expression in F(T)III, inferred above, could also well explain the negative correlation between this factor and the criterion normal-sociopathic (Cattell & Scheier, 1961).

At its positive pole, factor F(T)IV, *High Educated Self-Consciousness v. Low Self-Consciousness*, involves such characteristics as high shrewdness and imagination, alert control, and high self-sentiment. The factor is interpreted as a (presumably largely environmentally determined) pattern of keen self-awareness, possibly originating in a critical, competitive environment.

Factor F(T)V, *History of Inhibiting, Restraining Environment v. Laxness,* is primarily loaded by the first-order factor U.I. 17. The hypothesis is that F(T)V represents a broad personality source trait of restraint upon, and inhibition of, overt behaviour, however not related to any maladjustment or neurotic restrictiveness. (Note that factor U.I. 23, which matches the negative pole of Eysenck's neuroticism factor, loads positively on F(T)V.)

In F(T)VI, *Narcissistic Development v. Environmental Contact and Investment,* the central characteristic is a self-directed, both narcissistic and autistic, development lacking sound reality contact—either because of active negation or passive ignoration. Somehow F(T)VI comes close to what the clinician might consider as infantile or regressed self-assertion.

Factor F(T)VII, *High Tension to Achieve and Controlled Drive Tension Level v. Low Tension to Achieve,* represents a broad personality source trait of well-channelled and goal-directed motivation and drive tension.

III. THIRD-ORDER FACTORS IN OBJECTIVE PERSONALITY TESTS: METHOD OF ANALYSIS AND RESULTS

(a) *Method of analysis*

Higher-order factor analyses undertake a factorization of the matrix of intercorrelations among the factors at the next lower level of analysis. However, this correlation matrix can be determined by two different approaches: (1) by actually measuring, by batteries, the factors at the preceding stage of analysis and computing their intercorrelations, or (2) from the angular separation, between the factors rotated to the best possible simple structure in hyperspace, at the end of a lower-order factor analysis. The second approach has been chosen here because it has the advantage that the intercorrelations (in our case of the second-order factors) are free from systematic attenuation due to both test unreliability and imperfect validity of the factor measurement. On the other hand, such correlations are not free from random errors due to inexact discovery of simple structure, and therefore it becomes extremely important to rotate very thoroughly until no further improvement of hyperplane count can be demonstrated.

The intercorrelations of the second-order factors, on which this analysis is based, were obtained from three different studies as shown in Table 5.

In each of the three analyses all seven second-order factors, F(T)I–F(T)VII, have been identified. The second-order factor intercorrelations were computed for each study directly from the intercorrelations of the reference vectors* in the final oblique simple structure position; the resulting three 7 × 7 correlation matrices were then averaged, using Fisher's z'-transformation. The obtained pooled matrix of intercorrelations of the second-order factors represents the basic data of the present analysis; since this matrix has already been published in the article by Cattell *et al.* (1961) it need not be reproduced here.

---

* For the distinction between reference vectors and factors see, for example, Harman (1960).

The decision as to the number of third-order factors to be extracted was based upon application of Guttman's criterion (Guttman, 1954) which indicated three factors (see Table 6). Communalities for three principal axes factors have been estimated iteratively; 15 complete iterations gave stable h^2-values. With these communalities as diagonals, the second-order factor correlation matrix was factored for three principal axes factors. Table 7 shows this orthogonal factor matrix.

Table 5. *The three researches from which the intercorrelations of second-order factors were obtained*

		Method of analysis	
Research*	Subjects	First-order factors	Second-order factors
R 1	86 college undergraduates	Factor analysis of test intercorrelations	Analysis of the intercorrelations of primary factors
R 2	86 college undergraduates		
N 1	315 U.S. Navy Submarine School candidates	First-order factors were measured	

* In studies R 1 and R 2 the same 86 college students served as subjects; the two (only partly overlapping) test batteries were administered on two different occasions.

For a description of the three researches R 1, R 2, and N 1 the reader is referred to the article by Cattell, Knapp & Scheier (1961).

Table 6. *Latent roots of the matrix of second-order factor intercorrelations (with unities as diagonals)*

Order of latent root	Size of latent root
1	+1·59096168
2	+1·32596163
3	+1·13903331
4	+0·88407078
5	+0·82599782
6	+0·68432757
7	+0·54964723

Table 7. *Results of principal axes factor analysis*

	Principal axes factor		
	1	2	3
F(T)I	−0·26	+0·26	−0·38
F(T)II	+0·21	+0·44	+0·40
F(T)III	+0·68	−0·08	−0·24
F(T)IV	−0·20	−0·41	+0·39
F(T)V	−0·25	+0·00	−0·31
F(T)VI	+0·51	−0·10	−0·01
F(T)VII	−0·11	+0·17	+0·18

The orthogonal factor matrix was submitted to two oblique analytical rotation procedures, Pinzka & Saunders' (1954) Oblimax and Cattell & Muerle's (1960) Maxplane. The former took four over-all rotations for convergence, the latter seventeen. The solutions obtained from the analytical rotation routines were further rotated visually for maximizing simple structure. The post-Oblimax solution required three over-all oblique rotations, the post-Maxplane solution seven. Table 8 gives the two resulting final reference vector structure matrices,[*] Table 9 the obtained intercorrelations of the third-order factors.

Both solutions are equivalent in over-all simple structure, each yielding a hyperplane count of 48 % (± 0·10 hyperplane width). In deciding between the two alternative (but very similar) positions the post-Maxplane solution was regarded as preferable (*a*) because of its lower third-order factor intercorrelations, and (*b*) because of its simpler loading structure. (F(T) IV has three high loadings in the post-Oblimax solution, but only two in the post-Maxplane solution.) Only loadings equalling or above 0·20 will be regarded as 'significant'.

Table 8. *Final reference vector structure matrices*

		Post-Oblimax. Factor			Post-Maxplane. Factor		
	Second-order factors	A	B	C	A	B	C
F(T) I:	Tied Socialization	−0·03	+*0·46*	+0·04	−0·03	+*0·47*	−0·11
F(T) II:	Expansive Ego	+0·04	+0·02	+*0·59*	+0·05	+0·01	+*0·63*
F(T) III:	Temperamental Ardour	+*0·70*	+0·08	−0·02	+*0·70*	+0·07	−0·07
F(T) IV:	Educated Self-Consciousness	−*0·40*	−*0·56*	−*0·27*	−*0·40*	−*0·55*	−0·09
F(T) V:	History of Inhibiting, Restraining Environment	−0·08	+*0·23*	−0·17	−0·08	+*0·24*	−*0·25*
F(T) VI:	Narcissistic Development	+*0·44*	−0·09	+0·01	+*0·44*	−0·09	+0·03
F(T) VII:	Tension to Achieve	−0·16	−0·00	+*0·20*	−0·16	−0·00	+*0·22*

Table 9. *Intercorrelations of third-order factors*

	Post-Oblimax				Post-Maxplane		
	A	B	C		A	B	C
A	+1·00	−0·34	+0·03	A	+1·00	−0·35	+0·05
B	—	+1·00	−0·36	B	—	+1·00	−0·05
C	—	—	+1·00	C	—	—	+1·00

(*b*) *Interpretation of results*

It is suggested that the three obtained third-order factors be indexed as F(T) A, F(T) B, and F(T) C, respectively. Considering the psychological meaning of their second-order salients, they can be interpreted as follows:

F(T) A. This is the largest of the three factors in terms of second-order factor variance contribution. Its principal loading is on F(T) III + (*Temperamental Ardour*); in addition, the factor loads on F(T) IV − (*Low Self-Consciousness*) and F(T) VI + (*Narcissistic Development*). Thus factor F(T) A is characterized by high exuberance of behaviour, increased self-concern and self-directed attitudes, low shrewdness and alertness, and by lack of adequate internal integration and control. Psychologically, the most distinct feature of this pattern is a lack of maturity[†] and balanced control. In fact, practically all characteristics of F(T) A + can be accounted for by this interpretation. No direct results are as yet available on the age correlation of this factor—the indirect evidence, however, is in support of this hypothesis (see footnote †, this page).

---

[*] The reference vector structure, as a matter of fact, is proportional by columns to the factor pattern (see, for example, Harman, 1960).

[†] U.I. 21 + and U.I. 22 −, principal salients of F(T) III + F(T) IV −, respectively, both decline through the age period 9–15 years (Cattell, 1957*a*, p. 617).

The slight negative loading of F(T)VII (*Tension to Achieve*) on F(T)A is also in accordance with this interpretation.

In order to contrast factor F(T)A from the mature pattern of (goal-directed) ego expansiveness, the descriptive factor title *Immature Self-Centred Temperament v. Mature Environment-Oriented Stolidity* is suggested.

F(T)B. This factor loads on three second-order factors: F(T)I+ (*Tied Socialization*), F(T)IV− (*Low Self-Consciousness*), and F(T)V+ (*History of Inhibiting, Restraining Environment*). Again the psychological pattern is very clear and internally consistent; at its positive pole this factor is characterized by a readiness to accept cultural norms and demands, a lack of ego strength and alertness (both F(T)IV− and U.I. 28+ in F(T)I+), and by cautiousness and restraint (U.I. 17+ as primary marker of F(T)V+). From an interpretative point of view this factor apparently implies high internalization of social rules and standards, this being brought about on the cost of reduced self-determination and self-assertion (U.I. 19− in F(T)I+) and by a (voluntary or enforced) readiness to restrict overt behaviour in the pursuit of fulfilling what is set as the 'norm' or 'good standard'. The fact that F(T)IV, which has its principal loading on this factor F(T)B, correlates significantly negatively with the criterion normal-sociopathic (Cattell & Scheier, 1961) would also agree with this interpretation. These 'comentive' attributes in F(T)B+ gave lead to the suggested descriptive label *Restrained Acceptance v. Unrestrained Neglect of External Norms.*

F(T)C. It is for this last factor that the two rotational solutions give slightly divergent results. As a matter of fact, with few variables *no* rotation device can give completely certain simple structure. However, if we choose the post-Maxplane instead of the post-Oblimax position, on the grounds discussed above, factor F(T)C+ loads on the three second-order factors (FT)II+ (Expansive Ego), F(T)V− (Laxness), and F(T)VII+ (Tension to Achieve). Of these the loading of F(T)II is by far the highest.

On the basis of these results the leading qualities in F(T)C+ are high efficiency of performance, determined self-assertion, and voluntary control over one's environment (U.I. 16+ and U.I. 19+ in F(T)II+)—together with low inhibition of behaviour (F(T)V−) and increased, but controlled and goal-oriented drive tension (F(T)VII+). A pervasive efficiency and competence (both in terms of speed and quality of performance) can be inferred for F(T)C+ from the nature of the primary factors marking its salient second-order factors; in the context of high goal motivation and low inhibition this suggests an interpretation of factor F(T)C in terms of high primary ego-control and ego-strength, paired with good general ability (U.I. 1+ in F(T)II+). The factor title *High Self-Assertion v. Low Self-Assertion* is suggested under the hypothesis that the assertive and dynamic qualities form the 'nucleus' of this personality source trait, and the good performance level is a function of those attributes rather than just an associated phenomenon.

The obtained intercorrelations between the third-order personality dimensions (see Table 9) are in good agreement with the relationships to be predicted from their psychological interpretation. Both rotational solutions indicate a substantial negative correlation between F(T)A (Immature Self-Centred Temperament) and F(T)B (Restrained Acceptance of External Norms) which ties in well with the respective factor interpretations. In the preferred post-Maxplane position factor F(T)C is practically orthogonal to the remaining factors, and both solutions again agree on a correlation of virtually zero between F(T)A and F(T)C.

4. DISCUSSION AND SUMMARY

Only minimal theoretical comment will be made on the above three patterns. Owing mainly to uncertainties inevitable in rotating with few variables, factor analysts would do well patiently to await replication on new subject samples before proceeding to the potentially extensive theoretical implications of such findings. However, certain considerations about higher-order factors in general should be borne in mind. Cattell (1946) has suggested, as an inductive conclusion, that correlations and loadings can be expected to become smaller in successive higher-order analyses. On the other hand, it is possible that this occurs only in the first step from tests to primary factors. The increase in the unique variance left unaccounted for with increasing order of factorization may be due to nothing more than that *fewer* common factors are discoverable, through shrinking size of analysed correlation matrices, at higher than at lower orders.

The commonly accepted position of Burt (1940) that a hierarchy of factors exists can be accepted only if it could be demonstrated that the 'hierarchy' is not an 'artefact' of this inevitable reduction of the number of correlated variates at higher orders. The writers would like to suggest the alternative theory of a 'spherical network' which supposes that higher-order factors are influences contributing to the variance of lower-order factors, but admits circular 'feedback' of causal action in which 'higher' factors could in turn be affected by those one or two orders lower. According to this, higher-order factors need not be any more 'important' than lower-order factors, though the probability remains that they will be broader in their influence.*

It is possible that what the *clinician* sees, without the aid of measurement and computers, must be restricted to relatively few factors and that these perceptions will tend to be restricted to seeing the highest-order factors. Although we definitely did not start our studies with any predilections for psychoanalytic theory, it is a striking fact that the psychoanalytic descriptions of Ego, Id and Superego would fit very well the three major patterns found in this research. We are not ready to adopt this as our theoretical position at present, for several other possibilities need to be explored. They can best be investigated by, first, independent checks by other investigators on the third-order factor structure (using the Objective-Analytic Test Battery (Cattell, 1955*b*) to measure the primary personality factors) in new population samples, and by clinical and experimental research relating actual measures on these three factors to various life criteria and experimental influences.

REFERENCES

ALLPORT, G. W. & ODBERT, H. S. (1936). Trait-names, a psychological study. *Psychol. Monogr.* 47, no. 211, 1–171.

BROGDEN, H. E. (1940). A factor analysis of forty character tests. *Psychol. Monogr.* 234, 35–55.

---

* The ratio k/n, where k is the number of common factors and n the number of variables, increases in the personality area from 1/5 at the first-order (tests to primaries) to 3/7 at the third-order level.

Brunswik, E. (1956). *Perception and the Representative Design of Experiments.* Berkeley: University of California Press.

Burt, C. (1940). *The Factors of the Mind.* London: University of London Press.

Cattell, R. B. (1945). The principal trait clusters for describing personality. *Psychol. Bull.* **42**, 129–61.

Cattell, R. B. (1946). *The Description and Measurement of Personality.* Yonkers, N.Y.: World Book Co.

Cattell, R. B. (1955*a*). The chief invariant psychological and psycho-physiological functional unities found by *P*-technique. *J. Clin. Psychol.* **7**, 319–43.

Cattell, R. B. (1955*b*). *The Objective-Analytic (O-A) Personality Factor Batteries.* Champaign, Ill.: Inst. Personal Ability Testing. (Revised edition in preparation.)

Cattell, R. B. (1957*a*). *Personality and Motivation Structure and Measurement.* New York: World Book Co.

Cattell, R. B. (1957*b*). A universal index for psychological factors. *Psychologia,* **1**, 74–85.

Cattell, R. B. (1961). Clinical diagnosis and prognosis by objective personality factors. Paper read at 1961 *Annual APA Convention* in New York City.

Cattell, R. B. & Beloff, H. (1956). La structure factorielle de la personalité des enfants de onze ans a travers trois types d'épreuves. *Rev. Psychol. Appl.* **6**, 65–89.

Cattell, R. B., Knapp, R. R. & Scheier, I. H. (1961). Second-order personality factor structure in the objective test realm. *J. Cons. Psychol.* **25**, 345–52.

Cattell, R. B. & Muerle, J. L. (1960). The 'Maxplane' program for factor rotation to oblique simple structure. *Educ. Psychol. Meas.* **20**, 569–90.

Cattell, R. B. & Saunders, D. R. (1950). Interrelation and matching of personality factors from behaviour rating, questionnaire, and objective test data. *J. Soc. Psychol.* **31**, 243–60.

Cattell, R. B. & Saunders, D. R. (1955). Beitraege zur Faktorenanalyse der Persoenlichkeit. *Z. exp. angew. Psychol.* **2**, 325–57.

Cattell, R. B. & Scheier, I. H. (1958). The nature of anxiety: A review of thirteen multivariate analyses comprising 814 variables. *Psychol. Reports,* **4**, 351–88.

Cattell, R. B. & Scheier, I. H. (1960). Stimuli related to stress, neuroticism, excitation and anxiety response patterns: illustrating a new multivariate experimental design. *J. Abnorm. (Soc.) Psychol.* **60**, 195–204.

Cattell, R. B. & Scheier, I. H. (1961). *The Meaning and Measurement of Neuroticism and Anxiety.* New York: Ronald Press.

Cattell, R. B., Schiff, H. *et al.* (1953). *Psychiatric Screening of Flying Personnel: Prediction of Training Criteria by Objective Personality Factors and Development of the Seven Factor Personality Test.* Rep. No. 10, Contract no. AF 33 (038)–19569. Project no. 21–0202-0007, Randolph Field, Texas: USAF School of Aviation Medicine.

Cattell, R. B., Stice, G. F. & Kristy, N. F. (1957). A first approximation to nature–nurture ratios for eleven primary personality factors in objective tests. *J. Abnorm. (Soc.) Psychol.* **54**, 143–59.

Cattell, R. B. & Warburton, F. W. (1964). *Principles of Personality Measurement and a Compendium of Objective Tests.* Urbana: University of Illinois Press (in the Press).

Eysenck, H. J. (1947). *Dimensions of Personality.* London: Kegan Paul.

Eysenck, H. J. (1952*a*). Schizothymia-cyclothymia as a dimension of personality. II. Experimental. *J. Personal.* **30**, 345–84.

Eysenck, H. J. (1952*b*). *The Scientific Study of Personality.* London: Routledge and Kegan Paul.

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometr. Monogr.* **5**.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika,* **19**, 149–61.

Harman, H. H. (1960). *Modern Factor Analysis.* Chicago: University of Chicago Press.

Hundleby, J. D., Pawlik, K. & Cattell, R. B. (1964). *The First Twenty-one General Personality Dimensions in Objective, Behavioural Measurements.* Chicago: University of Chicago Press (in the Press).

Knapp, R. R. & Most, J. A. (1960). Personality correlates of marine corps helicopter pilot performance. *U.S.N. Med. Fld. Res. Lab. Rep.* No. 180109.1.3.

Mood, A. McF. (1950). *Introduction to the Theory of Statistics.* New York: McGraw-Hill.

PAWLIK, K. (1961). *Educational prediction from objective personality test dimensions*. Paper read at 1961 Annual APA Convention in New York City.

PINZKA, C. & SAUNDERS, D. R. (1954). Analytic rotation to simple structure. II. Extension to an oblique solution. *Research Bull. RB*-54-3. Princeton, N.J.: Educational Testing Service.

RYANS, D. G. (1938). An experimental attempt to analyze persistence. *J. Gen. Psychol.* **19**, 333-53.

THORNTON, G. R. (1939). A factor analysis of tests designed to measure persistence. *Psychol. Monogr.* **51**, 1-42.

THURSTONE, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.

THURSTONE, L. L. (1944). *A Factorial Study of Perception*. Chicago: University of Chicago Press.

(*Manuscript received* 14 *January* 1963)

Brit. J. Psychol. (1964), 55, 1, *pp.* 19–30
Printed in Great Britain

THE RELATIVE IMPORTANCE OF PROACTIVE INHIBITION AND DEGREE OF LEARNING IN RETENTION OF PAIRED ASSOCIATE ITEMS

BY PETER B. WARR
Department of Psychology, University of Sheffield

It has been reported that recall scores tend to decline with increasing numbers of lists learned. This has been attributed to proactive inhibition from the earlier lists. Two experiments are reported which test the hypothesis that this variation in recall score is more dependent on variations in mean degree of learning of items. This more economic explanation is supported by the findings, but analysis of the data indicates that a second factor plays a minor role in determining recall scores. It is suggested that this may be proactive inhibition, but that other interpretations are also possible.

It has been suggested by Underwood (1957) that an important learning variable affecting retention is the previous learning experience of subjects. Underwood collates the results of several earlier experiments to demonstrate that recall of verbal material is a function of the number of previous lists learned: the greater the number of lists learned, the lower is the recall score after 24 hours. This is interpreted to mean that the greater the number of previous lists learned the greater is the proactive inhibition. Proactive inhibition is in this way taken as a causal factor affecting recall performance.

It has, on the other hand, often been reported that subjects require fewer trials to criterion with increasing numbers of lists presented. This has, for instance, been shown for serial lists by Ward (1937) and Postman (1962), and for paired associate lists by Greenberg & Underwood (1950). It is possible that this increase in speed of learning determines at least part of the decline in recall performance with increasing numbers of lists learned. The experiments reported in this paper were conducted in order to test this hypothesis.

The experimental procedure employed followed, in outline, the paradigm discussed by Underwood (1957). Subjects learned a series of lists on successive days, with recall of each list being measured before the next list was learned. The learning task was to associate correctly paired adjectives. It should be noted at this point that retention of paired associates, as measured by a single relearning trial after 24 hours, is expected to be better than retention of a serial list, primarily because ordering of items is of greater importance in recall of a serial list. This implies that it may be misleading to infer the magnitude of the decline in recall performance after several lists have been learned from a composite curve presenting the results of both paired associate and serial learning. Of the results comprising the composite curve exhibited by Underwood (1957) only three are in fact of experiments employing solely paired associate lists. It is significant that these three experiments (Weiss & Margolius, 1954; Gibson, 1942; Williams, 1950) result in higher recall scores than the remaining experiments cited by Underwood. Although, therefore, we expect to find a decline in recall scores with an increase in the number of paired associate lists learned, we do not expect this decline to be as large as that indicated by Underwood's composite curve.

EXPERIMENT I

Experimental procedure

Materials

One hundred and forty-four two-syllable adjectives were selected from the sample of students' spoken vocabulary observed by Black & Ausherman (1955). One half of these adjectives occurred less than six times in this sample (of 288,152 word-symbols) and served as response terms in the present experiment. The other 72 adjectives occurred more than six times in Black & Ausherman's sample, and these served as stimulus terms. One half of the stimulus terms were 5-letter words and the other half were of 6 letters. The set of response terms was similarly constituted. Stimulus and response terms were paired in such a way that these differences in word-length were counterbalanced. Additional limitations on the selection and pairing of these adjectives were as follows: (1) No adjective which could also be used as a noun or a verb was selected (e.g. 'arctic'). (2) No term obviously similar to a previously selected adjective was employed (e.g. 'civil' and 'civic' were not both selected). (3) An attempt was made to avoid any 'expected' pairings of adjectives.

Six twelve-item lists were drawn from the 72 pairs of items. This was on a haphazard basis except that three pairs in each list were from each of the four pairing categories shown above. It is felt that the six lists so generated were of approximately equal difficulty.

Method and subjects

Each list was learned to a criterion of one correct trial by the anticipation method. The list was projected on a screen in a darkened room from slides arranged so that one order of presentation of the list was in each slide magazine. Nine magazines were filled for each list, so that nine different orders of presentation were possible. If more than nine trials were required, the tenth order of presentation was the same as the first, the eleventh as the second and so on. The interval between presentations of the list was 10 sec.

Each slide was presented for 3 sec. with an inter-slide interval of 1 sec. For each pair the stimulus item alone was first presented and subjects' anticipations were recorded by having each subject write a single anticipation on each page of an answer booklet. The stimulus item was then presented coupled with the response item.

Twenty undergradute subjects took part in the experiment and these were divided into five groups of four. This allowed five different arrangements of magazines. It also allowed observation of the response of each subject to ensure that the following conditions were met: nothing should be written while both the stimulus and the response terms were being displayed; each subject should indicate the correctness of each response by a tick or cross; each subject should note the number of each trial in the booklet as announced; each subject should leave the room on anticipating correctly all response terms on one trial. No difficulty was experienced by subjects in following these instructions.

The experimental design required subjects to be present at the laboratory at the same time on 7 successive days. A single list was learned to criterion on each day and recalled the following day. The recall measure was an additional anticipation trial. Three minutes after the recall of the previous list, learning of the current list commenced. In this way, six lists were learned and each was recalled after 24 hours. Each subject sat in the same position on each day and conditions were as nearly identical over days as was possible. Although subjects were aware that they were to recall a given list on the next day, no attempt was made by instructions to prevent rehearsal of a list outside the experimental situation. There is some doubt about the effect on retention of instructions designed to prevent rehearsal (e.g. Underwood & Keppel, 1962), and indeed about the effect of such instructions on rehearsal. It may be noted at this point that, in response to questions at the end of the experiment, six subjects admitted some rehearsal, but only of items from list 1. There is of course no way of checking the validity of this estimate.

Results and discussion

Learning performance

As had been anticipated, speed of learning increased with increasing number of lists learned. Learning speed can be measured in several ways, some of which may be considered here:

(*a*) *Trials to criterion*: subjects' performance on this traditional measure is shown in Table 1. A steady decline in mean trials to criterion is accompanied by decreasing intersubject variability as evidenced by the S.D. values.

(*b*) A somewhat more refined measure of learning speed is obtained by determining the *mean trial on which items of a list are first correctly associated.* This performance measure does not suffer from the disadvantage attached to the trials to criterion measure that a small number of exceptionally difficult items in the list may exert an undue influence. Table 1 also shows the results obtained by applying this second measure to the data.

Table 1. *Learning performance for successive lists in Experiment* 1 *described by three measures*

	List					
	1	2	3	4	5	6
Mean trials to criterion	11·40	10·20	8·80	7·75	6·95	6·85
S.D.	3·67	3·54	3·14	3·05	2·56	1·56
Mean trial on which first correct association occurs	5·55	4·80	4·35	4·20	3·66	3·70
S.D.	1·23	0·81	0·66	0·78	0·68	0·46
Mean number of correct anticipations per item	5·90	5·63	4·92	4·21	3·97	3·86
S.D.	1·79	1·74	1·37	1·17	1·37	0·91

(*c*) It is possible to divide the total learning period into a response learning stage and an associative stage (Underwood, Runquist & Schulz, 1959; Underwood & Postman 1960; Warr, 1963). In a paired associate learning task, response learning involves the acquisition of response terms in a given list as relevant to that list, whereas during the associative phase these reponses come to be associated with the prescribed stimulus term. Postman (1961, 1962) has suggested that the *duration of the response learning stage* can be estimated by the number of the trial on which the correct response is given anywhere in the list, and that the *length of the associative stage* can be assessed by the number of trials between response learning and the first correct association of responses. It will be seen that this procedure is in fact a refinement of measure (*b*) above, in that the mean total duration of these two phases is equivalent to the mean trial on which items of a list are correctly associated (i.e. measure (*b*)). Table 2 exhibits a decline in both stages with increasing numbers of lists learned.

(*d*) A measure that is particularly suited to the analysis which follows is obtained by observing the *mean number of correct anticipations per item*. This measure is in fact determined not only by the number of trials to criterion but also by how early in learning a subject starts correctly to associate items. For instance, subjects requiring

247439

fewer trials to criterion will not necessarily make fewer correct anticipations per item, since it is possible that they will start to make correct anticipations earlier than slower learners. The mean number of correct anticipations per item for successive lists is shown in Table 1.

Table 2. *Duration of response-learning and associative stages in learning successive lists in Experiment* 1

	List					
	1	2	3	4	5	6
Mean length of response-learning stage	4·40	4·20	3·98	3·85	3·37	3·52
Mean length of associative stage	1·15	0·60	0·37	0·35	0·29	0·18

Mean speed of learning increases over successive lists according to each of these four measures. This has important consequences in terms of mean degree of learning of items. Degree of learning is here defined operationally as the number of times that an item is correctly associated during learning. No assumptions are made in this definition about the 'strength' of associations, resistance to extinction, etc. It follows from this operational definition that measure (*d*) above is in fact also a measure of degree of learning. Mean degree of learning of items can be seen from Table 1 to decline over the six lists.

Recall performance

The recall performance of the group 24 hours after learning each list is shown fully in Table 3. This table exhibits the finding that mean *C* score (items correct) declines steadily over the six lists. A similar decline in mean *S* score is also shown. *S* score is a measure of the number of substitute responses made. These are responses legitimate in that they occur in the current list but are attached to an inappropriate stimulus term. Overt interference from previous lists (*I* score) is also shown in Table 3, according to the list from which such overt interference occurred. It is significant that, although mean *I* score increases over the six lists, it on no occasion reaches a high enough value to account for the decline in *G* score. *St* score (recall of stimulus terms) and *X* score (emission of extraneous, non-experimental items) show no consistent changes throughout the experiment. *M* score, which is also shown in Table 3, is the total number of responses made, being therefore the total of *C*, *S*, *I*, *St*, and *X* scores.

The *percentage* of items correctly recalled after 24 hours is shown in Fig. 1 (the curve labelled 'Experiment 1'). A decline from 82·5 to 60·8% occurs over the six lists. This decline is very similar to the decline in *paired associate* recall performance in the data collected by Underwood (1957). It is less marked than the decline shown in Underwood's composite curve, but as indicated above, recall of paired associate lists is expected to be better than recall of serial lists. Although, therefore, the observed decline in *C* score may provide some support for the proactive inhibition hypothesis, proactive inhibition cannot always be as important a factor as Underwood suggests.

Further analysis of the data, however, indicates that it is invalid to attribute this decline wholly to proactive inhibition. It has already been noted that mean degree of learning of items (as measured by number of correct anticipations) declines with increasing numbers of lists learned. The parallel decline in mean degree of learning

Table 3. *Recall performance for successive lists in Experiment* 1

	List					
	1	2	3	4	5	6
Mean *C* score	9·90	8·70	8·60	8·25	8·05	7·30
Mean *S* score	0·75	0·45	0·25	0·35	0·30	0·20
Mean *I* score:						
List minus 1	—	0·15	0·30	0·50	0·15	0·50
List minus 2	—	—	0·00	0·10	0·00	0·10
List minus 3	—	—	—	0·00	0·00	0·10
List minus 4	—	—	—	—	0·00	0·00
List minus 5	—	—	—	—	—	0·10
Mean *St* score	0·05	0·05	0·15	0·10	0·05	0·00
Mean *X* score	0·00	0·05	0·05	0·00	0·00	0·15
Mean *M* score	10·70	9·40	9·35	9·30	8·55	8·45

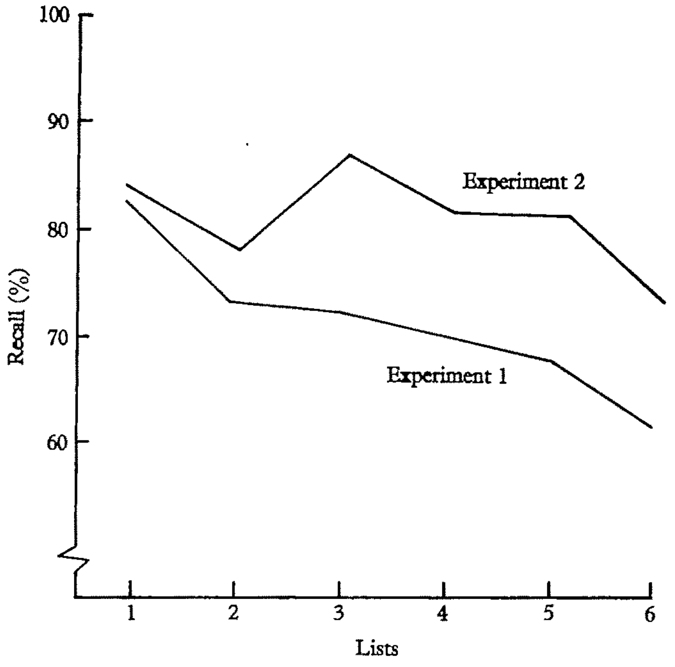


Fig. 1. Mean percentage of items recalled correctly after 24 hours from each of the six lists in Expts. 1 and 2.

and *C* score suggests that degree of learning may be a factor affecting the number of items recalled. It is clear, however, that a rather more direct relationship has to be established before this conclusion can validly be drawn.

If it is possible to show that the recall of individual items is dependent on the degree of learning of those items, we can argue that the decline in mean degree of learning is in fact producing at least part of the decline in *C* score.

The most straightforward way of investigating the relationship between degree of learning of each individual item and the probability of recall of that item is to relate number of correct anticipations to the probability of recall. If this is done it is possible to pool the data for all subjects and produce a graph with 'Number of correct anticipations' and 'Probability of recall' as the axes. This method has been used by Underwood (1954). It is, however, possible that an artefact is introduced by the fact that not all subjects contribute to each probability value, since some subjects need fewer trials to criterion than others. The former subjects will, therefore, only contribute to the probability values for the smaller number of correct anticipations.

Runquist (1957) has used a second method of combining number of correct associations data for a group of subjects. He avoids the possible artefact considered above by ranking each item for each subject according to the maximum number of correct anticipations made by each subject. Thus, the item receiving the maximum number of correct anticipations for each subject is ranked '1' irrespective of the actual number of correct anticipations. It then becomes possible to combine the data for each rank so that each subject contributes to each combined score. This procedure is also open to a methodological objection: it does not follow that the intervals between ranks are the same for all subjects. The method permits a difference between any number of correct anticipations to be equivalent to the difference between any number of ranks. For instance, if the maximum number of correct anticipations of an item by a subject is 12, the item receiving 12 correct anticipations is ranked '1'. Consider the case of two subjects for both of whom the item ranked '1' has been correctly anticipated 12 times. Suppose that for one of these subjects the item ranked '2' has been correctly anticipated 11 times, and that for the other subject the item ranked '2' has been correctly anticipated 8 times. It is unlikely that to combine the data of these two subjects for rank 2 is a legitimate step.

This difficulty may be avoided by a slightly more precise ranking procedure. The number of correct anticipations received by each item for a given subject may be expressed as a percentage of the maximum number of correct anticipations made by that subject. In this way rank 1 is always equivalent to 100%; rank 2 in the hypothetical case above is equivalent to 92% (i.e. $\frac{11}{12} \times 100$) and 67% (i.e. $\frac{8}{12} \times 100$) for the two subjects. But whereas this method avoids the objections levelled against the two previous measures, it is itself not wholly satisfactory. This is because each value is dependent on the maximum number of correct anticipations for each subject. In this way a given percentage value for different subjects can be a function of several widely differing numbers of correct anticipations.

Since each of these three methods of relating degree of learning to probability of recall is somewhat unsatisfactory in one respect, the present data were analysed according to each of the procedures to find whether the relationships were systematically distorted by the potential artefacts discussed. The potential artefact is of a different kind in each case and each method should yield a different relationship if the artefact is of importance. The results of this analysis suggest that these possible artefacts are of little significance, since the obtained relationship between degree of learning and probability of recall after 24 hours is the same irrespective of the procedure used.

It is found that the probability of recall, after 24 hours, of individual items from each of the six lists is directly dependent upon degree of learning, as measured by number of correct anticipations during learning. Fig. 2 shows this relationship for each list on the basis of the first procedure outlined above. The probability of correct recall measure ($(P(C))$) is of the relative frequency type, each value being the number of items recalled after each number of correct anticipations during learning divided by the total number of items receiving that number of correct anticipations. The categories along the abscissa have been grouped in pairs in order to increase the reliability of the $P(C)$ measures obtained in this manner. It should be noted that whereas the $P(C)$ values for smaller numbers of correct anticipations are reliable, those for higher numbers of correct anticipations are based on only small samples and cannot be taken to be reliable. For example, only five items were correctly anticipated eleven and twelve times during learning of list 3. The $P(C)$ value after

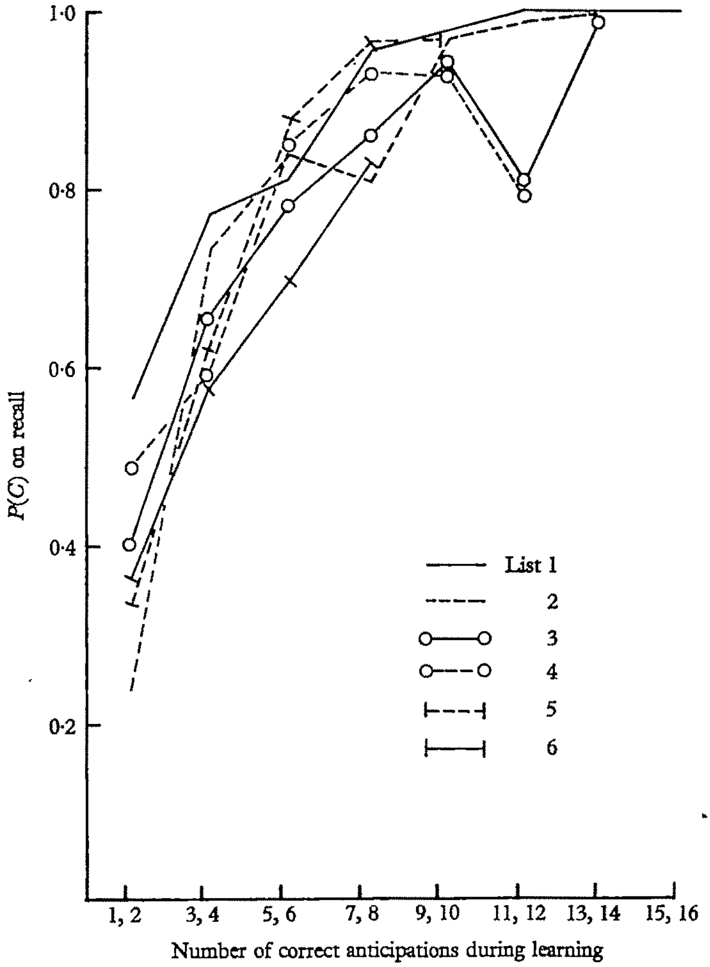


Fig. 2. The relationship between degree of learning of individual items in Expt. 1 (measured by number of correct anticipations) and the probability of recall ($P(C)$) after 24 hours. Each curve indicates this relationship for a given list.

eleven and twelve correct anticipations of items in this list can be seen from Fig. 2 to be 0·80. This indicates that four of the five items in this category were recalled after 24 hours; this proportion is based on too small a sample to be reliable.

The argument so far may be summarized as follows:

(1) Mean degree of learning of items in each list has been found to decline with increasing numbers of lists learned.

(2) The probability of recall of individual items after 24 hours has been shown to depend directly on the degree of learning of those items.

(3) The finding that mean recall score declines with increasing numbers of lists learned is at least partly explicable by a combination of (1) and (2).

An explanation of subjects' recall performance under these conditions in terms of decline in mean degree of learning of items would of course be simpler than an explanation invoking the additional concept, proactive inhibition. It is felt that this simpler explanation is to a large extent valid. It is, however, possible that an additional factor is contributing to the decline in C score with increasing numbers of lists learned. Yet if the argument thus far is valid, any second factor can only play a minor role in the explanation of subjects' recall performance.

If any factor apart from degree of learning is depressing subjects' recall of later lists, we should expect a given degree of learning to be associated with a lower probability of recall with each additional list learned. The probability of recall for each degree of learning of each list is shown in Fig. 2; it is interesting that the list 6 curve in Fig. 2 is consistently below the list 1 curve. This indicates that a given degree of learning (as measured by number of correct anticipations) does not raise the probability of recall of items to as high a level when five previous lists have been learned as when no previous lists have been learned.

In order to investigate whether a reduction in $P(C)$ for each degree of learning is brought about by each additional list learned, it was necessary to discover whether the following two conditions could be met: (*a*) the curves in Fig. 2 could be shown to be significantly different; (*b*) the overall rank-order of the curves could be shown to follow the order in which the lists were learned, so that the earlier the list, the greater the $P(C)$ of items for each degree of learning. It was found that the first of these two conditions could, in fact, be met; a Friedman two-way analysis of variance (combining values above seven correct anticipations) indicated that the curves were significantly different ($\chi_r^2 = 11{\cdot}34$, $P < 0{\cdot}05$). The second requirement (that there is a significant tendency for the $P(C)$ of items for each degree of learning to decline with each additional list learned) was tested by computing the rank correlation (ρ) between the overall rank order of the curves in Fig. 2 in terms of $P(C)$ values and the order in which the lists were learned. A significant positive correlation indicates that a given degree of learning is associated with a lower probability of recall after 24 hours with each additional list learned. This is the minimum requirement for an explanation in terms of proactive inhibition.

It was found that $\rho = +0{\cdot}66$. With $N = 6$ this value of ρ is not statistically significant. It cannot, therefore, be argued that the present results provide evidence for the influence of proactive inhibition on the recall of later lists. The results are best explained in terms of the decline in mean degree of learning of items with increasing numbers of lists learned. On the other hand, the obtained ρ of 0·66 is certainly suggestive, so that the influence of an additional factor ought not to be dogmatically ruled out.

EXPERIMENT 2

This experiment was conducted in order to investigate further the effects of systematic variation of degree of learning on recall of successive lists. It has been shown above that speed of learning increases as subjects learn a number of lists on successive days. When learning is terminated after the first correct anticipation of the whole list, mean degree of learning, as here defined, declines for successive lists. It is, therefore, to be expected that if the number of learning trials is held constant for a series of lists, mean degree of learning will *increase* with increasing numbers of lists learned. On the basis of the theoretical framework outlined in this paper, we should expect that recall scores after 24 hours should fail to decline over lists in this condition. We should also expect to find the same relationship as in Expt. 1 between degree of learning of individual items and the probability of recall of those items.

Experimental procedure

Degree of learning was manipulated by terminating learning after 12 trials on each day. The materials of Expt. 1 were also used in this experiment and recall was again measured by a single anticipation trial 24 hours after learning. Twenty-one undergraduate subjects learned the lists in the same order as used in the previous experiment. Apart from the fact that number of trials per list was held constant, experimental conditions were identical with those obtaining in Expt. 1.

Table 4. *Learning performance for successive lists in Experiment 2*

	List					
	1	2	3	4	5	6
Mean correct anticipations per item	5·77	6·13	7·67	7·57	8·30	8·31

Results and discussion

Learning performance

Speed of learning was again found to increase under these conditions. Of the four measures of learning speed considered earlier only the fourth is applicable to this design. The increase in learning speed according to this measure is shown in Table 4. Since the number of trials per list was held constant, a greater number of correct anticipations per item indicates faster learning. Degree of learning has been operationally defined in terms of number of correct anticipations; it can be seen from Table 4 that mean degree of learning increases with increasing numbers of lists learned under these conditions.

Recall performance

The recall performance of the group 24 hours after learning each list is shown fully in Table 5. It can be seen that the mean number of items recalled correctly after 24 hours (C score) does not show the marked decline observed in C score in Expt. 1. This is of course partly due to the fact that subjects had not learned all the items in, for instance, list 1. The question of retention clearly does not arise for unlearned items. C score has, therefore, also been related to original learning in Table 5. The measure of original learning employed is the number of items correctly associated on the last learning trial (i.e. trial 12). *C* score as a percentage of original learning remains relatively constant over the six lists (see also Fig. 1). As in Expt. 1, overt interference from previous lists (*I* score) is infrequent. Subjects' 'learning to learn' is also evidenced by a decline in the number of substitute responses (*S* score) with increasing numbers of lists learned.

The relationship between degree of learning of individual items and probability of recall was investigated in the same way as in the previous experiment. This relationship is exhibited in Fig. 3 for each of the six lists. Probability of recall of items is again directly dependent on degree of learning, as measured by number of correct anticipations during learning. It may, therefore, be concluded that the fact that recall scores after 24 hours do not decline with increasing numbers of lists is attributable to the increase in mean degree of learning over those lists.

Table 5. *Recall performance for successive lists in Experiment* 2

	List					
	1	2	3	4	5	6
Mean *C* score	8·29	7·57	9·33	8·95	9·05	8·24
C score as % of original learning	83·65	77·94	85·22	80·34	80·17	71·49
Mean *S* score	0·81	0·62	0·14	0·24	0·38	0·33
Mean *I* score:						
List minus 1	—	0·24	0·19	0·10	0·05	0·24
List minus 2	—	—	0·05	0·05	0·00	0·05
List minus 3	—	—	—	0·19	0·05	0·00
List minus 4	—	—	—	—	0·10	0·00
List minus 5	—	—	—	—	—	0·00
Mean *St* score	0·10	0·00	0·00	0·05	0·10	0·00
Mean *X* score	0·14	0·14	0·00	0·05	0·10	0·00
Mean *M* score	9·34	8·57	9·71	9·63	9·83	8·86

It would appear, however, that a secondary factor is affecting these results. This can be shown in two ways: first, if degree of learning of individual items were the only determinant of the subject's recall performance, *C* score as a percentage of original learning should increase as degree of learning increases. That it does not increase can be seen from Fig. 1. And secondly, the relationship between the $P(C)$ curves for the six lists (Fig. 3) is systematically and significantly non-random. The test applied to the results of Expt. 1 was also applied here. A Friedman analysis of variance revealed that the curves are significantly different ($\chi_r^2 = 18{\cdot}28$; $P < 0{\cdot}01$). The influence of additional lists learned on the $P(C)$ value for each degree of learning was again assessed by calculating the rank correlation between the overall rank order of the curves in Fig. 3 in terms of $P(C)$ values and the order in which the lists were learned. In the present case $\rho = +0{\cdot}83$ ($P = 0{\cdot}05$). In other words, a given degree of learning (as measured by the number of correct anticipations) is significantly associated with a lower $P(C)$ after 24 hours with increasing numbers of lists learned.

It must, therefore, be concluded that, although the observed increase in mean degree of learning of items is the major factor determining the present results, a second factor is also present. It is possible that this second factor is proactive inhibition. It may, however, be noted that other interpretations of this finding are possible. For instance, the degree of extra-experimental rehearsal may decline with each additional list learned; or subjects may become increasingly proficient in the techniques required for learning by the anticipation method. But the present results leave little doubt that this second factor, whatever its nature, is playing only a minor part in determining subjects' recall performance.

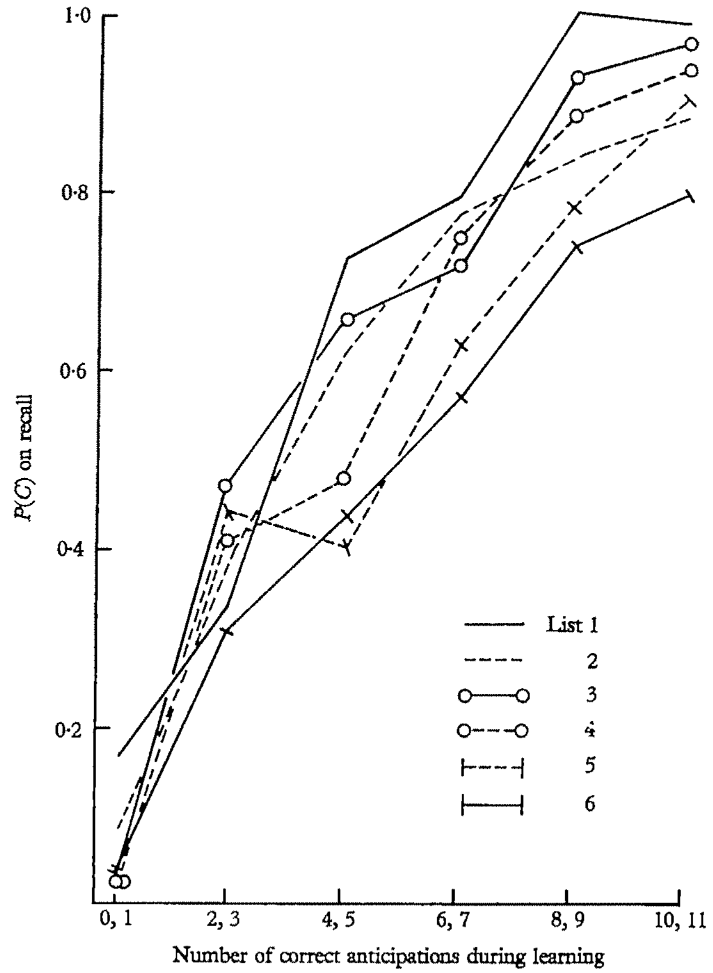


Fig. 3. The relationship between degree of learning of individual items in Expt. 2 (measured by number of correct anticipations) and the probability of recall ($P(C)$) after 24 hours. Each curve indicates this relationship for a given list.

Conclusions

The following conclusions appear to follow from this analysis of the results of the two experiments described here:

(1) When lists of paired associate items are learned to a criterion of one correct trial by the anticipation method, recall scores after 24 hours tend to decline with increasing numbers of lists learned. The decline in recall scores of six paired associate

lists in the present case was from 82·5 to 60·8%: this is in accord with the paired associate data presented by Underwood (1957) but not as great a decline as appears to take place when a number of *serial* lists are learned and recalled.

(2) When lists of paired associate items are learned for a constant number of trials, recall scores after 24 hours do not decline with increasing numbers of lists learned.

(3) The probability of recall of individual items after 24 hours is dependent on degree of learning of those items, as measured by the number of correct anticipations.

(4) Variations in the mean degree of learning of items account for the major part of the observed variation in the recall scores obtained throughout a series of lists. This is the case for lists learned to criterion and for lists learned for a fixed number of trials.

(5) A minor part of the variation in recall scores obtained throughout a series of lists is not attributable to differences in degree of learning. This minor part of the variation may perhaps be termed proactive inhibition, although other interpretations are possible.

(6) It would, therefore, seem that Underwood (1957) overestimates the influence of proactive inhibition on the recall of later lists learned.

This work was supported by a grant from D.S.I.R. to the Department of Psychology, University of Sheffield for research into training procedures. The writer is grateful to Prof. H. Kay and Mr J. K. Clarkson for their constructive criticism during the preparation of this paper.

REFERENCES

BLACK, J. W. & AUSHERMAN, M. (1955). *Vocabulary of College Students in Classroom Speeches*. Ohio State University–U.S. Naval School of Aviation Medicine, Pensacola, Florida.

GIBSON, E. J. (1942). Intra-list generalization as a factor in verbal learning. *J. Exp. Psychol.* **30**, 195–200.

GREENBERG, R. & UNDERWOOD, B. J. (1950). Retention as a function of stage of practice. *J. Exp. Psychol.* **40**, 452–7.

POSTMAN, L. (1961). Extra-experimental interference and the retention of words. *J. Exp. Psychol.* **61**, 97–110.

POSTMAN, L. (1962). The temporal course of proactive inhibition for serial lists. *J. Exp. Psychol.* **63**, 361–9.

RUNQUIST, W. N. (1957). Retention of verbal associates as a function of strength. *J. Exp. Psychol.* **54**, 369–75.

UNDERWOOD, B. J. (1954). Speed of learning and amount retained. *Psychol. Bull.* **51**, 272–86.

UNDERWOOD, B. J. (1957). Interference and forgetting. *Psychol. Rev.* **64**, 49–60.

UNDERWOOD, B. J. & KEPPEL, G. (1962). An evaluation of two problems of method in the study of retention. *Amer. J. Psychol.* **75**, 1–17.

UNDERWOOD, B. J. & POSTMAN, L. (1960). Extraexperimental sources of interference in forgetting. *Psychol. Rev.* **67**, 73–95.

UNDERWOOD, B. J., RUNQUIST, W. N. & SCHULZ, R. W. (1959). Response learning in paired associate lists as a function of intralist similarity. *J. Exp. Psychol.* **58**, 70–8.

WARD, L. B. (1937). Reminiscence and rote learning. *Psychol. Monogr.* **49**, no. 4 (Whole no. 220).

WARR, P. B. (1963). The effect of repeated presentation of items on paired associate learning. *Quart. J. Exp. Psychol.* **15**, 262–72.

WEISS, W. & MARGOLIUS, G. (1954). The effect of context stimuli on learning and retention. *J. Exp. Psychol.* **48**, 318–22.

WILLIAMS, M. (1950). The effects of experimentally induced needs on retention. *J. Exp. Psychol.* **40**, 139–51.

(*Manuscript received* 30 *November* 1962)

Brit. J. Psychol. (1964), 55, 1, *pp.* 31–37
Printed in Great Britain

THE EFFECTS OF THE VARIATION OF TWO PHYSICAL PARAMETERS ON EMPIRICAL DIFFICULTY IN A PERCEPTUAL MAZE TEST

BY ALICK ELITHORN,* DAVID JONES,† MYFANWY KERR*
AND DAVID LEE†‡

National Hospitals for Nervous Diseases and the Royal Free Hospital, London

An improved version of a perceptual maze test, which has been shown to be sensitive to frontal and temporal lobe damage, is described. The present test possesses a fully binary structure and parameters describing the task are now more easily definable. An experiment is reported in which the effect of varying two of these parameters, size and density of pattern, was examined. In order to minimize differential practice effects, a 36 × 36 random Latin square design was used, 36 normal subjects being tested with 36 patterns. The results show that the effect of the variation of these two parameters on the difficulty experienced by the subject in solving the task is reilably pronounced. A third parameter 'pattern' is briefly discussed.

INTRODUCTION

A majority of intelligence tests are of a verbal nature and the value of such tests as tools for the experimental study of intelligent behaviour is limited by the difficulty of controlling cultural and educational bias. 'Non-verbal' tests in general are less subject to this disadvantage but involve the discrimination of relationships between geometrical figures or patterns whose parameters are not easily definable. For this reason the relative complexity of individual items in such tests cannot be expressed logically but is evaluated in terms of the relative difficulty of the items, determined empirically.

Theories of intellectual function include both single and multiple factor theories. In practice, single factor theories postulate at least one specific, or group factor in relation to any particular test performance and some current workers tend to regard the concept of a single factor for intelligence—the ability to educe relationships—as an artefact of the statistical approach to the study of thinking. The work of several psychologists, (Halstead, 1947; Teuber, 1950) as well as the studies of the present authors (Elithorn, 1955; Elithorn, Kerr & Mott, 1960), has recently underlined the importance of perceptual integrative mechanisms as correlates of the ability to order data intelligently. This approach appears to stem from the work of Lashley (1938) who expressed the view that perceptual skills are phylogenetically fundamental components of intelligent behaviour.

These considerations suggest that an intellectual task which is perceptual in nature, which possesses a structure which can be analysed mathematically and which is as free as possible from cultural bias, may prove a useful tool for the study of intellectual skills. Such a task in the form of an intelligence test has been described in two previous papers (Elithorn, 1955; Elithorn *et al.* 1960). The early observations showed that this test has internal consistency and suggested that an ability which

---

* Member of the Medical Research Council External Staff.
† Supported by a grant from the Medical Research Council.
‡ Working also at the University of London Computer Unit.

contributes to performance in other intelligence tests is reliably sampled. These conclusions and the observations that the test is sensitive to frontal and temporal lobe damage have been confirmed by a more recent study undertaken in America by Prof. Arthur Benton (Benton, Elithorn, Fogel & Kerr, 1963). The present experiment makes a first step in investigating systematically the sources of difficulty in this material by studying the effects of varying two parameters, the size of the lattice through which a path was to be traced and the number of target dots which are superimposed on the lattice (see Fig. 1).

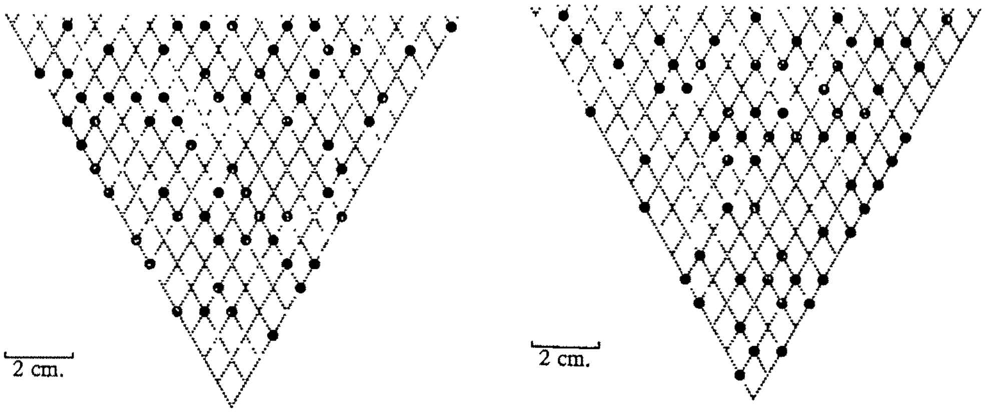


Fig. 1. Two maze patterns, both 16 row, 40 % saturation. The one on the left (No. 128) is comparatively easy, being solved by 26 subjects. The one on the right (No. 134) is more difficult and was solved by only 8 subjects. The solution numbers are 11 and 12 respectively.

In the previous version of the test the patterns were rectangular in form. This meant not only that the pattern structure was not fully binary in that the edges provided only single choice points, but in addition, the variable 'shape' was not controlled since the relative length of the sides was allowed to vary. Thirdly, the position of the pattern dots was not determined by objective criteria but reflected the designers' idea of what would constitute a difficult pattern; there was also a specific limitation in that each pattern contained only one set of dots which provided a solution pathway. With the present design it was decided not to determine the positioning of the dots by subjective criteria. This avoids the limitation that only specific types of patterning can be examined and means that significant patterning factors other than those being tested may be detected. The placing of the dots has therefore been randomized, using random number tables.

MATERIALS AND METHODS

The maze patterns which constitute the test material are triangular lattices represented by dotted lines. At a variable number of intersections (excluding the vertex) larger black target dots are superimposed (Fig. 1). The subject is required to find a pathway along the dotted lines from the vertex to the top of the maze which passes through the greatest possible number of target dots; there are two rules: the pathway must lie completely on the lattice lines and must not cut across the intervening spaces; at each intersection the pathway may fork only right or left and must not double back. The maximum number of dots obtainable in each pattern is printed at the bottom right-hand corner of each pattern.

Two parameters of the maze, size and saturation, have been varied systematically. These parameters may be defined:

(1) 'n': the number of horizontal rows in the lattice excluding the vertex (size).

(2) 's': the ratio of number of dots to total number of lattice points (saturation).

Each of these two variables was controlled at three levels giving nine types of test item ($n = 12, 16, 20$ and $s = 0{\cdot}2, 0{\cdot}3$ and $0{\cdot}4$).

To minimize learning effects a 36 × 36 randomized Latin square design was used, 36 subjects being given 36 maze patterns, 4 of each of the 9 types defined above. These were presented in a random order with the restriction that no one pattern was presented in the same ordinal position in a set to more than one subject. The 36 subjects were in-patients at the Royal Free Hospital, N.W. 3, or, in a few cases, members of staff at the same hospital. All were between the ages of 14 and 55 and none had any known cerebral damage or psychological impairment.

The instructions given to each subject were based on those used in previous tests, quoted above, and in all cases demonstration patterns were used to ensure that the subject had fully understood the rules and object of the task before starting the test.

Each subject was tested individually with one maze pattern at a time, the time taken for each pattern being recorded. He was told to work as quickly as possible as there was a time limit, but was asked not to draw in a pathway until he had found the correct one. (The time limit was 1 min. though the subject was not specifically informed of this.) It was stressed that the time allowed might be inadequate for some patterns and that if he was unable to find the solution before it expired, he would be allowed a further 15 sec. in which to draw in the best pathway he had found. He was also told that the patterns were not arranged in any order of difficulty.

RESULTS

Figs. 2*a*, *b* show the distribution of successes between subjects and patterns (a success is recorded if a subject finds a correct pathway within the time limit allowed). In Fig. 2*c* the number of successes recorded for the patterns is related to both their size and saturation.

Although differences in size and saturation clearly correlate with observed difficulty, the horizontal scatter within rows shows that the remaining variable, which can be broadly defined as the arrangement of the dots or 'patterning', is an important determinant of item difficulty. The effect of this variable is most marked with the 16, 0·4 and 20, 0·2 mazes which show a wide range of scores. The smaller mazes are evidently too easy and the larger mazes with high saturation too difficult to reveal patterning as an important source of variance.

The significance of the two parameters being examined (n and s) has been tested by a 2-way cross classification analysis of variance. The results are given in Table 1. The *interaction* between the main factors just fails to reach the 0·05 level of significance ($F = 2{\cdot}63$) and the main factors have therefore been tested against the new residual R' which includes the variance due to their interaction. Both the main factors are significant at above the 0·1 % level of confidence. Since the interaction factor is marginally significant, a breakdown analysis of variance has been undertaken. The results are shown in Table 2.

This table shows that saturation is non-significant as a main factor for the smallest size of patterns but significant at the 0·05 and 0·01 levels for the middle and largest sizes respectively. This confirms the suspicion that saturation as a determinant of difficulty is most important in relation to certain pattern sizes. Table 2(*b*), however, shows that size is significant at all three levels of saturation, confirming its significance as a main factor in the main analysis.

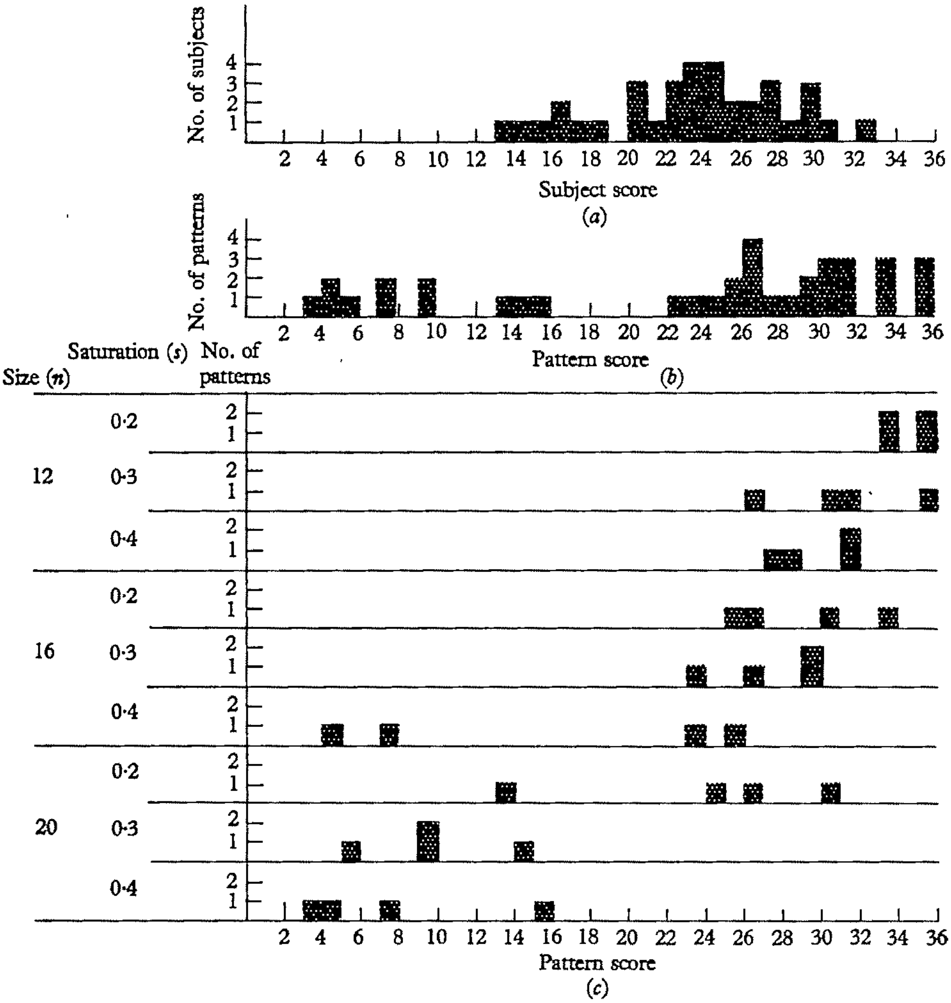


Fig. 2. Distribution of successes.

The relationship of size and saturation to experimental difficulty, though satisfactorily demonstrated in this experiment, cannot be defined in detail since the factors were tested only at three levels, which were selected arbitrarily. It is possible that a different choice of values might give different results. Different functional relationships might be expected to exist between experimental difficulty and each of these variables. As a first hypothesis it seems likely that difficulty will vary exponentially with size, but will increase monotonically with saturation to around the 50 % level and thereafter decrease again monotonically. Although the present data are compatible with these hypotheses, they are quite inadequate for any exact test.

Table 1

Source	d.f.	Sum of squares	Mean square	F-ratio (factor/R)	P	F-ratio (factor/R')	P
Size (n)	2	1952	976	34·28	$P \ll 0·001$	28·31	$P \ll 0·001$
Saturation (s)	2	798	399	14·01	$P < 0·001$	11·57	$P < 0·001$
Size × saturation ($n \times s$)	4	300	75	2·63	$P = 0·05+\epsilon$ (very small)	—	—
Residual (R)	27	768·75	28·47	—	—	—	—
Total	35	3,818·75	—	—	—	—	—
$R' = R+(n \times s)$	31	1,068·75	34·48	—	—	—	—

Table 2*a*

		Size ($n = 12$)				Size ($n = 16$)				Size ($n = 20$)			
Source	d.f.	Sum of squares	Mean square	F	P	Sum of squares	Mean square	F	P	Sum of squares	Mean square	F	P
Saturation	2	48	24	375	NS	441	220·5	4·69	$< 0·05$	608	304	9·5	$< 0·01$
Residual	9	58	6·4	—	—	423	47	—	—	288	32	—	—
	11	106	—	—	—	864	—	—	—	896	—	—	—

Table 2(*b*)

		Saturation ($s = 0·02$)				Saturation ($s = 0·03$)				Saturation ($s = 0·04$)			
Source	d.f.	S.S.	M.S.	F	P	S.S.	M.S.	F	P	S.S.	M.S.	F	P
Size	2	231	115·5	5·09	$< 0·05$	1020	510	39·8	$< 0·001$	1001	500·5	10·0	$< 0·01$
Residual	9	204	22·7	—	—	115	12·8	—	—	450	50	—	—
	11	435	—	—	—	1135	—	—	—	1451	—	—	—

The present experiment was designed to assess the practicability of relating the concept of subjective difficulty to certain parameters of the task rather than to provide the basis of a practical test. However, some conclusions on this latter point can tentatively be drawn. The sample of subjects used contained individuals from all social classes, as defined by the Registrar General, including some of university calibre. On the whole, the 36 items provided good discrimination between subjects. Scores were fairly normally distributed from 12 (the score of the worst subject) to 33 (the score of the best) (Fig. 2*a*). With regard to the patterns themselves, no one pattern proved so difficult as to be completely insoluble while three were solved by everyone. This implies that the simplest patterns are too easy for normal subjects. A further point worthy of comment is that the distribution of scores for individual items shows a dichotomy, the item ranked 11th in order of difficulty being solved by 16 subjects, while item 12 was solved by 23 subjects. This break in the distribution cannot be related to too abrupt a change in the parameters systematically varied (size and saturation) since at the critical level at which the split occurs 2 of the (16, 0·4) patterns and 3 of the (20, 0·2) fall into the easy group and the rest of these two sets into the difficult group. Presumably, this break should be related to a patterning factor. Since this has not been systematically controlled but allowed to vary in a random way, it is unlikely that there is any large step increase in pattern complexity. It is possible, however, as suggested in an earlier paper, that a systematic increase in complexity leads at a certain level to a change in the strategies used by the subject. If this is the explanation, then it is a little surprising that the effect should be so well defined in the group results. A comparable effect has, however, been shown in the Holyoke study of perceptual span (Kaufman, Lord, Reese & Volkmann, 1949). In this work it was shown that the recognition of number could be achieved rapidly for samples of seven items or less by a perceptual process—'subiting'—but that for higher numbers accuracy could only be achieved by the slower strategy of 'counting'. The findings in the present experiments therefore support the suggestion made previously that when patterns reach a certain degree of complexity 'perceptual scanning' gives place to a more rigorous strategy of systematic search.

DISCUSSION

The present experiment has established the importance of pattern size and saturation as determinants of subjective difficulty. It has also, however, confirmed the existence of a source of difficulty derived from the arrangement or distribution of the target dots at the interstices of the matrix. The relative significance of this patterning factor could not be assessed with the present factorial design since the position of the dots was determined randomly and the effect is confounded with error variance. It is possible, however, to demonstrate its importance by evaluating for the whole group the within-group value of χ^2. This value ($\chi^2 = 84{\cdot}06$; d.f. 27) corresponds to a probability of considerably less than 0·001 that the difference in difficulty for each level of size and saturation is an artifact due to sampling variance. The effect is demonstrated most clearly by the four patterns of middle size and highest saturation which ranked 3, 6, 13 and 15 in order of difficulty and were solved by 5, 8, 24 and 26 subjects respectively ($\chi^2 = 35{\cdot}3$; d.f. 3; $P < 0{\cdot}001$). It is worth noting also that since in this subset there are $\binom{152}{61}$ patterns, or approximately $1{\cdot}91 \times 10^{43}$, it is unlikely that in a random sample of four two will be from among the most difficult and two from among the least difficult. It is not unreasonable therefore to conclude that the range of difficulty attributable to patterning in this subset is wide and that a range of items matching the abilities of a normal sample of the population could be selected from this subset alone. It seems to be indicated therefore that the next step should be a detailed analysis of the patterning factor, possibly with the factors of size and saturation held constant at a uniform level.

Patterning parameters can be related both to the characteristics of the solution pathway and to the topological relationships which exist between the sets of dots on the solution pathway and the remaining dots. In preparing the earlier rectangular version of this test (Elithorn, 1955), it appeared that difficult solutions possessed all or some of the following characteristics:

(1) The ratio of the number of pathways giving an optimum score (m) to the number of pathways giving a score of one less ($m-1$) was low. (In the rectangular version the solution pathways passed through only one series of dots, but with random selection the 'best' pathway may not be unique.)

(2) The ratio of the number of inflexions on the solution pathway to the maximum possible on a pathway was high.

(3) There was a large number of choice points in which an immediate gain had to be given up for a later gain.

(4) The optimum paths lay in areas which seemed perceptually less attractive than other areas (e.g. horizontal groupings give an appearance of target clustering which is in fact irrelevant to the task solution.)

In a preliminary analysis based on the present results, there is evidence to support the existence of the first of these sources of difficulty: that is, patterns with a low ratio:

$$\frac{\text{no. of pathways with maximum score } (m)}{\text{no. of pathways with score } (m-1)}$$

were found to be more difficult on the average than patterns of the same 'm' value with higher ratios. However, this effect is confounded with other sources of patterning difficulty and further analysis on the present data seems hardly justified. A mathematical formulation which will permit a systematic analysis of these and similar psychological causes of difficulty is being developed (Buckingham, Elithorn, Lee & Nixon, 1963).

We are grateful to the staff and patients of the Royal Free Hospital for their help with the testing, and to Dr R. A. Buckingham and Mr W. L. B. Nixon for helpful discussion on the problem of pattern difficulty.

REFERENCES

BENTON, A., ELITHORN, A., FOGEL, M. & KERR, M. (1963). A perceptual maze test sensitive to brain damage. *J. Neurol. Neurosurg. & Psychiat.* **26**, *No.* 6, 540–4.

BUCKINGHAM, R. A., ELITHORN, A., LEE, D. N. & NIXON, W. L. B. (1963). A mathematical model of a perceptual maze test. *Nature, Lond.*, **199**, 667–8.

ELITHORN, A. (1955). A perceptual maze test sensitive to brain damage: A preliminary report. *J. Neurol. Psychiat.* **18**, 287–92.

ELITHORN, A., KERR, M. & MOTT, J. (1960). A group version of a perceptual maze test. *Brit. J. Psychol.* **51**, 19–26.

HALSTEAD, W. C. (1947). *Brain and Intelligence: A Quantitative Study of the Frontal Lobes.* Chicago: University of Chicago Press.

KAUFMAN, E. L., LORD, M. W., REESE, T. W. & VOLKMANN, J. (1949). The Discrimination of Visual Number. *Amer. J. Psychol.* **62**, 498–525.

LASHLEY, K. S. (1938). The mechanism of vision. 2. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **18**, 123–93.

TEUBER, H. L. (1950). Neuropsychology. In *Recent Advances in Diagnostic Psychological Testing*. Springfield Ill.: Thomas.

(*Manuscript received* 2 *January* 1963)

Brit. J. Psychol. (1964), 55, 1, *pp.* 39–48
Printed in Great Britain

SHAPE DISCRIMINATION BY RATS: SQUARES AND RECTANGLES

BY N. S. SUTHERLAND AND A. E. CARR
Institute of Experimental Psychology, Oxford University

Sixteen rats were trained to discriminate between a square and a rectangle: both shapes were presented in different orientations to different animals. All animals were trained in succession with positive shape only, positive and negative shapes presented simultaneously, positive and negative shapes presented successively; finally they were given tests with 18 transfer shapes presented singly. Although the rats showed a strong preference for the square early in training, they later responded more accurately to the rectangle presented on its own than to the square. There was no tendency to transfer from a rectangle in one orientation to rectangles in other orientations, and indeed when a horizontal rectangle was substituted for a vertical or vice versa negative transfer resulted. The rats (unlike octopuses) showed complete transfer to outline shapes. The transfer test results are discussed in relation to recent physiological discoveries about retinal receptive fields.

I. INTRODUCTION

It has recently been shown that there are some striking similarities between the ways in which octopuses and rats discriminate between visually presented shapes (Sutherland & Carr, 1962). It is of interest to establish how far shapes are discriminated in the same way by two such different species: if *differences* in the ways shapes are classified can be discovered, then it should be possible to correlate them with differences in the visual environment or neuro-anatomy of the two species and thus to throw light on the mechanisms underlying shape discrimination.

The present experiment replicates with rats an earlier experiment performed on octopuses (Sutherland, 1960). Different rats were trained to discriminate between squares and rectangles in different orientations, and were then given transfer tests with a variety of further shapes. The theoretical rationale and implications of the experiment will be discussed below.

A second purpose of the experiment was to confirm two findings of a previous experiment on rats. Sutherland, Carr & Mackintosh (1962) trained rats to discriminate between an open and a closed shape presented simultaneously; the animals exhibited a strong initial preference for more open shapes, and when subsequently performing in a successive situation their responses to the open shapes were more accurate than their responses to the closed shapes. By an 'open' shape is meant one which is spread out (e.g. a cross), by a 'closed' shape one which is compact (e.g. a circle). Earlier work suggests that both rats (Sutherland & Carr, 1962) and octopuses (Sutherland, 1963*b*) analyse shapes along a dimension which has been labelled 'the open–closed dimension' because it corresponds approximately to the loose definition of 'open' and 'closed' given above. However, it is not possible to say whereabouts a given shape will lie on this dimension for a given animal merely by inspecting the shape: this can only be determined by running transfer tests and discovering from animals' responses how they actually order the shapes along the dimension. Both octopuses and rats have been found to treat rectangles as more open shapes than squares or diamonds and to treat an upright cross as more open than an 'X'. In what follows, when one shape is said to be more 'open' or more 'closed' than another, this refers to the experimentally determined positions of the shapes on the dimension for rats (or octopuses). A fuller discussion of the terms and further evidence that this is one of the primary dimensions along which rats and octopuses discriminate shapes will be found in Sutherland (1963*a*). Since rectangles are more open for rats than squares, it is of interest to discover whether rats in fact discriminate between these two shapes in terms of their position along the open–closed dimension and the present series of transfer tests provides evidence on this.

The following terms will be used with the definitions given here. (1) *Square*: any square in any orientation. (2) *Normal square*: a square with base horizontal. (3) *Diamond*: a square with sides running at 45° to horizontal and vertical. (4), (5): *Vertical rectangle, horizontal rectangle*: these terms are self-explanatory. (6) *Oblique rectangle*: a rectangle with sides running at 45° to horizontal and vertical.

II. METHODS

Subjects

The subjects were 16 male hooded rats from the colony maintained at the Oxford University Institute of Experimental Psychology. They were about 4 months old at the start of the experiment.

Apparatus and procedure

The apparatus and training procedures used have been fully described elsewhere (Sutherland, 1961; Sutherland *et al.* 1962). Rats were placed in a start box with a transparent door through which they could view two shapes attached to a vertical wall 12 in. away. (At some stages of the experiment only one shape was exposed at a time, see below.) After a 5 sec. delay, the door was raised and rats were allowed to run to one or other of the shapes, which were placed to left and right of a partition projecting 3·5 in. from the far wall into the choice chamber. All shapes were cut from $\frac{1}{4}$ in. white Perspex. If a rat approached to within 3·5 in. of the negative shape a bell rang and the animal was removed from the apparatus. If it approached the positive shape, it was allowed to drink for 9 sec. from a milk tube projecting through the centre of the shape. A modified correction procedure was used throughout: if animals ran to the wrong side, they were re-run without changing the position of the shapes until they chose correctly. Only errors on the first run of each trial were counted.

The experiment was conducted in four stages. (1) Animals were first given 100 trials pretraining to run to the positive shape exposed on its own with a milk tube on the other side of the apparatus. (2) All subjects were then given ten trials a day of simultaneous training with both shapes exposed, until they ran for 2 consecutive days without making an error. After each rat reached this criterion, it was given ten trials every third day, until all rats had done so. This occurred on the 45th day of training. (3) Animals were then retrained to discriminate between the same shapes presented singly, with a milk tube but no shape on the opposite side. When the positive shape was presented, they had to continue to run to it; when the negative shape was shown, they had to run to the milk tube exposed on the opposite side. Twelve trials were given a day at this stage, six with the positive and six with the negative shape, and retraining continued for 22 days. (4) Finally, animals were given a series of transfer trials each of which was preceded by one retraining trial with one or other of the original shapes presented singly. Each animal was given ten transfer tests with each of 18 different transfer shapes: 12 retraining trials and 12 transfer trials were given a day. The order in which transfer shapes were shown was randomized. On transfer trials the bell was turned off, and animals were allowed 9 sec. access to the milk whether they ran to the transfer shape or to the opposite side on which a milk tube on its own was exposed. The transfer shapes used will be described in the section on results. Unless otherwise stated, they were of approximately the same areas as the original training shapes.

Experimental design

Each rat was trained to discriminate between a square in one orientation and a rectangle in one orientation. There were two possible orientations for the square (normal square and diamond) and four for the rectangles (vertical, horizontal, and two oblique rectangles at 90° to one another) making eight combinations, and two rats were trained on each combination: of each such pair of animals, one was trained with square positive, one with rectangle positive. The rectangles used in training and retraining measured 10 × 2 cm., and the squares 5 × 5 cm.

III. RESULTS

Training

Over the first 30 trials of training, all animals for whom the rectangle was the positive shape scored below chance making only 32 % correct responses. Subjects being trained with square positive averaged exactly 50 % over the same period, and the difference in the means of the two scores was significant at better than the 0·01 level of confidence ($t = 3{\cdot}4$, d.f. = 14). This indicates that at the outset of training animals had a preference for the square. From day 5 onwards there were no consistent differences between the performance of animals being trained in opposite directions.

Table 1. *Percentage correct over days* 13–45 *of training*

Training shape	Vertical rectangle	Horizontal rectangle	Oblique rectangle
Normal square	95	79	85
Diamond	94	90	79

Table 1 shows the percentage of correct responses made by animals being trained on different shapes between days 13 and 45: exactly the same rank order for the six groups is obtained if we use the day on which criterion was reached as an index of performance. The difference between the scores made by the four animals trained to discriminate between a vertical rectangle and either square and the four trained to discriminate between a diamond and an oblique rectangle is significant at the 0·01 level ($t = 4{\cdot}1$, d.f. = 6). None of the other differences is significant.

Retraining

The results of retraining with single shapes are very similar to those obtained in an earlier experiment (Sutherland *et al.* 1962). At first animals made most of their mistakes by running to the negative shape exposed on its own: thus over the first 4 days of retraining animals ran *to* the positive shape on 89 % of trials, whereas they *ran away* from the negative shape on only 48 % of trials. Over the last four days of retraining they scored 88 % correct with the positive shape and 74 % with the negative. The same tendency was still apparent over the additional retraining trials given intermixed with transfer tests: over this period animals averaged 88 % correct responses when the original positive shape was exposed on its own and only 65 % correct with the original negative shape.

A further effect of considerable interest appeared in the course of retraining. Animals tended to do better when the rectangle was presented on its own than when the square was presented. Thus, over the 180 retraining trials intermixed with transfer tests, animals being trained with rectangle positive were correct when it was shown on 90 % of trials, whereas animals trained with square positive were correct when the square was shown on 86 % of trials. Animals trained with rectangle negative were correct on 77 % of trials with the negative shape, animals trained with square negative on only 53 % of trials. This difference was evaluated in the following way. For each animal the number of correct responses made with the negative shape was subtracted from the number made with the positive: these difference scores were averaged separately for animals trained with rectangle positive and for animals trained with square positive. The difference scores were much larger for the former group than for the latter, and the difference in the two means was significant at the 0·001 level of confidence ($t = 4{\cdot}6$, d.f. = 15). This indicates that animals trained with rectangle positive tended to do relatively better on their positive shape and relatively worse on their negative than animals trained with square positive; i.e. there was a highly significant tendency for animals to perform better when the rectangle was shown than when the square was shown provided the effects of better performance on positive than on negative shapes are partialled out.

Table 2. *Results of transfer tests*

Transfer from square			Transfer from rectangle			
	Transfer shape	Performance		Transfer shape	Group trained on	Performance
	Original square	69		Original rectangle		83
1*	Large square	61	13**	Large rectangle		80
2*	Small square	58	14	Small rectangle		58
3*	Outline square	71	15**	Outline rectangle		81
4*	45° rotated square	61	16	Horizontal rectangle	Vertical rectangle	8
5**	22·5° rotated square	68	17	Oblique rectangle	Vertical rectangle	19
6	Equilateral triangle base down	65	18	Vertical rectangle	Horizontal rectangle	38
7**	Equilateral triangle apex down	64	19	Oblique rectangle	Horizontal rectangle	51
8*	Circle	66	20	Vertical rectangle	Oblique rectangle	43
9**	X	64	21	Horizontal rectangle	Oblique rectangle	43
10	+	56	22	Opposite oblique rectangle	Oblique rectangle	43
11**	45° rotated, outline square	63				
12**	Outline circle	66				

* Transfer significant at 0·05 level of confidence; ** transfer significant at 0·01 level of confidence.

Transfer tests

The results of transfer tests are set out in Table 2. For the transfer shapes shown on the left of the table, the percentage of trials on which animals treated them as a square was computed, i.e. the percentage of runs made *to* each shape by rats trained with square positive, and the percentage of runs *away from* each transfer shape made by rats trained with square negative. The right-hand columns show the results for the remaining transfer shapes (various rectangles), and transfer is there measured as the percentage of times each shape was treated as a rectangle. For some of the shapes in the right-hand columns (nos. 16–22) the results are broken down according to which of the original rectangles was used in training. The figures for the first shape in each column represent performance during retraining trials with the original shapes and have already received comment. The better performance with the original rectangle than with the original square must, however, be borne in mind when interpreting the results of the transfer tests.

The significance of the transfer test results was evaluated in the following way. For tests 1–15 the significance of the difference between the scores of animals trained with square positive and those trained with rectangle positive was assessed by a non-parametric test for unpaired replicates (Wilcoxon, 1949). If for a given transfer shape the difference between the scores of animals originally trained in opposite directions is significant, this implies that their performances on the shape was influenced by their training and hence that significant transfer is occurring. The results of this test are given in Table 2. Unfortunately, it is not possible to apply this test to the results for transfer shapes 16–22 in Table 2 because of the small number of animals in each group. However, significant transfer was clearly occurring in tests 16 and 17. In test 16, the two animals trained with square positive each ran to the horizontal rectangle 10 times out of 10 presentations, whereas the two trained with vertical rectangle positive ran to it 3 times and no times. In test 17, the two animals trained with square positive ran to one or other of the oblique rectangles 19 and 17 times out of 20 presentations and the two trained with vertical rectangles positive only 2 and 7 times. When it was necessary to compare performance on one transfer shape with performance on another, a non-parametric test for the difference between means with paired replicates was conducted. In describing the detailed results, it is worth comparing them with previous findings obtained from octopuses. Arabic numbers in brackets refer to the transfer shapes given in Table 2.

(i) Animals transferred better to larger shapes (twice the linear dimensions of the originals) (1, 13) than to smaller (one half the linear dimensions) (2, 14). The difference was significant only for the large and small rectangles ($P < 0{\cdot}02$). This result has previously been obtained with rats trained on other discriminations (Fields, 1931, 1932; Sutherland, 1961) and also holds for octopuses trained on this and other discriminations (Sutherland & Carr, 1963).

(ii) The outline shapes used had outlines of 0·75 cm. breadth. Transfer to outline versions of the training shapes was complete (3, 15). Transfer from the square to two other outline shapes (11, 12) was as good as transfer to filled-in versions of the same shapes (4, 8). When octopuses had been trained to discriminate between squares and rectangles, they showed complete transfer to an outline rectangle but incomplete transfer to an outline square and much less transfer to an outline circle than to a filled-in circle (Sutherland, 1960).

(iii) There was good but incomplete transfer from the original square to several other compact shapes (4–8) of approximately the same area: this finding was also obtained with octopuses.

(iv) Rats trained with vertical rectangles treated a horizontal rectangle as a square, not as a rectangle (16); similarly animals trained with horizontal rectangles treated a vertical rectangle as a square (18). So did octopuses.

(v) Animals trained with vertical rectangles treated an oblique rectangle as similar to a square, whereas animals trained with horizontal rectangles showed no transfer to oblique rectangles. In the octopus experiment, there was a slight tendency for animals trained on vertical and horizontal rectangles to treat oblique rectangles as a square.

(vi) There was no transfer from oblique rectangles to either horizontal or vertical rectangles (20, 21) (there was very little in the octopus experiment).

(vii) There was no transfer from one oblique rectangle to its mirror image (22) (i.e. an oblique rectangle rotated through 90° from the orientation of the original). Octopuses showed practically perfect transfer between these two shapes.

(viii) An 'X' was treated as more similar to a square or diamond than was an upright cross ($P < 0{\cdot}02$). Although these shapes were not shown in transfer tests to octopuses trained to discriminate a square from a rectangle, it has been found in other experiments that octopuses treat an 'X' as a more closed shape than an upright cross (e.g. Sutherland, 1963*b*).

In Table 2 only the results with transfer rectangles in different orientations have been broken down according to the shape used in training. Results for all other transfer shapes were analysed in the same way, but they are not published here since groups trained on different shapes varied little in the way they responded to transfer shapes. The only consistent variation found was in the treatment of the upright and inverted triangle by groups trained on normal square and diamond. The former group treated the upright triangle as a normal square on 76 % of trials whereas the latter treated it as a diamond on 54 % of trials; the corresponding figures for the inverted triangle were 55 and 70 %. The upright triangle was more nearly equivalent to a normal square than to a diamond whereas the inverted one was more nearly equivalent to a diamond than to a normal square. Since the upright triangle has the same base line as a square, and the inverted one approximately the same base line as a diamond, this suggests that in the present experiment there was some tendency for rats to discriminate the shapes in terms of differences in their base lines. This was not the only feature of shapes discriminated, however, since most of the results do not accord with this hypothesis. Thus, animals trained on vertical rectangles showed less tendency to treat an upright cross as a rectangle than did animals trained with horizontal rectangles, although the lower part of the upright cross is identical with the lower part of the vertical rectangle. Again, animals trained on a vertical rectangle and a diamond showed just as much transfer from square to horizontal rectangle as did animals trained with vertical rectangle and normal square, although the base line of the horizontal rectangle bears no resemblance to that of the diamond.

IV. DISCUSSION

Preferences and learning

At the beginning of simultaneous training animals exhibited a preference for the square; during successive training they performed better with the rectangle than with the square. These findings are puzzling for several reasons.

(1) In two earlier experiments (Sutherland *et al.* 1962; Sutherland, 1964), it was found that at the outset of simultaneous training rats showed a strong preference for the more open shape of the two on which they were being trained. In the present experiment the less open shape is preferred. Moreover, in another experiment (Sutherland & Carr, 1962) a measure of preference was obtained after extensive training and it was found that an oblique rectangle was preferred to a diamond.

Although the earlier experiments strongly suggest that one of the factors influencing rats' preferences for shapes is their 'openness', the present results indicate that at least early in training it is not the only factor operating. It is possible that early in training the rectangles have some sort of negative valence.

(2) Granted that initially the rats preferred a square to a rectangle, the question arises of why, although the group being trained with rectangle *negative* ran significantly more often to the square than to the rectangle during the first 30 trials of training, animals being trained with square *positive* did not run to it more frequently than chance expectation over the same period. It seems likely that two factors are at work—the preference for the square, and a tendency to explore unfamiliar shapes. The effect of pre-training animals to run to a shape was in fact to make them avoid that shape when it was presented with another shape. On the very first day of simultaneous training this tendency was apparent even for animals trained with square positive: of these subjects five chose the rectangle more often than the square and the other three chose each shape equally often. The two effects—preference for the square and tendency to run to a new shape—work in the same direction for animals being trained with square negative; hence these animals ran to the square on 68 % of trials over the first 3 days. The effects work in opposite directions for animals trained with square positive and appear to have cancelled one another out so that these animals ran equally often to each shape over the first 3 days of training.

(3) The third puzzle in the present results is that, whereas in an earlier experiment (Sutherland & Carr, 1962) during successive training animals performed better with the shape they initially preferred, in the present experiment animals performed better with the less-preferred shape. The earlier results suggested that open shapes are more attention-compelling (have a higher arousal value) than closed shapes: hence they are preferred early in training, and performance during successive training is better on such shapes than on less attention-compelling shapes. The present results confirm that performance with shapes exposed singly is better with the more open shape (the rectangle), but we are still left with the puzzle that this shape was the less frequently chosen early in training. The problem cannot be resolved at the moment, but it seems clear that initial preferences are to some extent determined by some factor other than the degree of openness of the shapes.

One other possible explanation of the better performance in successive discrimination with rectangles than with squares should be mentioned. It may have been more difficult to discriminate the square presented alone from the milk tube on the opposite side of the apparatus than to discriminate the rectangle from the milk tube. This explanation seems implausible because of the disparity in size between the training shapes, whose areas were 25 cm.^{2} (squares) and 20 cm.^{2} (rectangles), and the milk tube with an area of about 0·05 cm.^{2}.

Shape-analysing mechanisms

The fact that animals transferred to shapes of different sizes from the originals is important in that it shows that discrimination was not based on the small areal differences between the training shapes. It also shows that discrimination was not based on an analysis of any one absolute dimension of the shapes.

We have already seen that some of the results of transfer tests suggest that animals' responses were partly determined by similarities in the lower halves of the shapes, but that this is not enough to explain all the transfer test data. It is possible that the differences found in initial training were also due to this factor. Thus the contours at the base of the oblique rectangles are orientated in exactly the same way as those at the base of the diamond, and the long horizontal contour at the base of the horizontal rectangle is similar to that at the base of the square and the animals trained on these two pairs were those which performed worst over days 13–45 of initial training (see Table 1).

The finding that rats do not discriminate less well between a diamond and an oblique rectangle than between a normal square and a horizontal rectangle contra-indicates a theory suggested by Dodwell (1957) to account for shape discrimination in rats (cf. also Sutherland, 1960).

The remaining findings can be interpreted in terms of recent discoveries made by Hubel & Wiesel (1962) about receptive fields connected to individual neurones in the striate cortex of cats. They found that the firing of single cells in the striate cortex could be modified by stimulation of large but well-defined areas of the retina (the cell's receptive field). Some such cells responded maximally to a thin rectangle projected on the centre of their retinal fields, others to a sharp boundary between light and dark. A given cell responded maximally to a stimulus in a particular orientation on the retina, and little or no response occurred if the stimulus was rotated through a small angle. Different cells responded maximally to stimuli in different orientations, and there was no evidence that cells responsive to any one orientation were more common than cells responsive to other orientations. There is anatomical evidence which suggests that not dissimilar receptive fields may exist in octopuses (Young, 1960) and in view of gross similarities in the neuro-anatomy of the visual system of cats and rats, it seems probable that such receptive fields also exist in rats. It is of course possible, and indeed probable, that the actual distribution of types of field will vary from species to species and some types may be present in some species and absent in others.

Sutherland (1963*a*) has recently shown that, if we postulate the existence of this type of coding in the visual systems of octopuses and rats, then much of the behavioural data on shape recognition in these two species can be explained. In octopuses it is probable on both behavioural and neuro-anatomical grounds (Sutherland, 1963*a*) that there are more receptive fields for detecting horizontally or vertically orientated stimuli than for detecting obliquely orientated stimuli. The findings of the present experiment can be explained if we make the following assumptions: (i) in the rat, receptive fields responding to boundaries are more predominant than in the octopus; (ii) the predominance of vertical and horizontal fields over oblique fields exists in the rat but is less marked than in the octopus; (iii) one operation performed on the output from striate cells connected to such fields is to count whether cells responsive to boundaries in a particular orientation fire more strongly than cells responsive to any other orientation; (iv) a second operation is to count the total number of cells firing. Each assumption will now be discussed in relation to the evidence.

(i) The rat was found to transfer much better to outline shapes than the octopus. Clearly if we take an outline of a square or diamond far more cells responding to thin segments of shapes will be fired than when a filled-in shape is presented. However, there will be little change in the number of boundary-detecting cells fired. Hence it seems likely that there are relatively more of the latter type of cell in the rat than in the octopus.

(ii) Both rat and octopus having been trained to discriminate a vertical rectangle from a square subsequently treat a horizontal rectangle like a square, whereas animals trained with square and horizontal rectangle treat a vertical rectangle as a square. This suggests that the original discrimination may have been performed in terms of the relative number of cells fired having vertically (or horizontally) orientated receptive fields. If this were so, then, for an animal trained on a square and vertical rectangle, a horizontal rectangle would resemble an exaggerated square—like the original square, it has more horizontal contour than the vertical rectangle and less vertical contour. Octopuses trained to discriminate between an oblique rectangle and a square subsequently treated an oblique rectangle at 90° to the training rectangle as equivalent to the original rectangle. Rats show a slight tendency to treat a rotated oblique rectangle in the same way as they had been trained to respond to the original *square*, though transfer was considerably less than that shown by rats trained with the vertical (or horizontal) rectangle and tested with a horizontal (or vertical) rectangle. The results make it quite clear that rats can discriminate between two oblique rectangles, but suggest that they discriminate less well between two oblique rectangles, than between a horizontal and vertical rectangle. Sutherland (1963*c*) found that cats discriminate as well between two oblique rectangles at 90° to one another as between a horizontal and vertical rectangle, and suggested that this may be connected with the fact that in the cat's striate cortex there appear to be as many cells responding to oblique retinal stimuli as there are cells responding to horizontal and vertical stimuli. The present results suggest that, in the rat, cells with vertical and horizontal receptive fields may predominate but that this predominance is less marked than in the octopus.

(iii) In the previous paragraph we have already used the assumption that, where striate cells responsive to stimuli in one retinal orientation fire more strongly than other cells, this may be detected and form the basis of discrimination. This would also explain the absence of significant transfer from a rectangle in one orientation to rectangles rotated 45° from that orientation. In addition, shapes *not* having contours running predominantly in one orientation are all treated as similar to a square (namely, rotated squares, circle, triangle, upright cross, and 'X').

(iv) The evidence that rats may also count the *total* number of cells fired by a given shape comes mainly from previous work on the way in which rats discriminate between open and closed shapes (Sutherland & Carr, 1962; Sutherland, 1963*a*). In the present experiment there was no transfer from one rectangle to another although any rectangle would fire more neurones with the type of receptive field envisaged than would a square. It would therefore seem that the present discrimination was achieved mainly by counting the predominance of firing of cells with a *given* orientation rather than simply the total firing. However, there was significantly more transfer from square to 'X' than from square to upright cross and we would expect this if some account is taken of the total number of cells fired, since the 'X', having oblique contours, should fire fewer cells than the upright cross, which has both horizontal and vertical contours, and hence of the two shapes the 'X' should be treated as being more similar to the original square.

It should be added that although in the case of octopuses and cats there appears to be a correlation between the relative number of neurones with oblique fields present and the animals' ability to discriminate between two obliques, it is by no means certain that the one phenomenon is the direct cause of the other. Thus it may be that, in animals which confuse two oblique rectangles, cells responding to boundaries orientated at 45° and 135° to the horizontal are at some further stage of the system connected to a common cell. It seems probable that such a mechanism for confusing two obliques would only be evolved in animals which for environmental reasons did not require to make fine discriminations between oblique contours. Hence such animals might also in the course of evolution come to have fewer cells detecting oblique orientations than detecting horizontal and vertical. Thus, we might expect the relative number of cells responding to stimuli in oblique orientations to correlate with a species' ability to discriminate between 90° rotated oblique rectangles but we cannot infer that scarcity of such cells is the direct cause of the confusion between two obliques.

This work forms part of a project on 'Stimulus analysing mechanisms': the project is financed by the American Office of Naval Research (contract N 62558–2453) and the Nuffield Foundation.

REFERENCES

DODWELL, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.* **48**, 221–9.

FIELDS, P. E. (1931). Contributions to visual figure discrimination in the white rat. II. *J. Comp. Physiol. Psychol.* **11**, 349–67.

FIELDS, P. E. (1932). Studies in concept formations. I. The development of the concept of triangularity by the white rat. *Comp. Psychol. Monogr.* **9**, 221–9.

HUBEL, H. & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–54.

SUTHERLAND, N. S. (1960). Visual discrimination of shape by *Octopus*: squares and rectangles. *J. Comp. Physiol. Psychol.* **53**, 95–103.

SUTHERLAND, N. S. (1961). Visual discrimination of horizontal and vertical rectangles by rats on a new discrimination training apparatus. *Quart. J. Exp. Psychol.* **13**, 117–21.

SUTHERLAND, N. S. (1963*a*). Shape discrimination and receptive fields. *Nature, Lond.*, **197**, 118–22.

SUTHERLAND, N. S. (1963*b*). The discrimination of stationary shapes by *Octopus*: further open and closed forms. *Amer. J. Psychol.* **76**, 177–90.

SUTHERLAND, N. S. (1963*c*). The cat's ability to discriminate oblique rectangles. *Science*, **139**, 209–10.

SUTHERLAND, N. S. (1964). A further experiment on the discrimination of open and closed forms by rats. *Quart. J. Exp. Psychol.* (in the Press).

SUTHERLAND, N. S. & CARR, A. E. (1962). Visual discrimination of open and closed shapes by rats. II. Transfer tests. *Quart. J. Exp. Psychol.* **14**, 129–39.

SUTHERLAND, N. S. & CARR, A. E. (1963). The visual discrimination of shape by *Octopus*: the effects of stimulus size. *Quart. J. Exp. Psychol.* **15**, 225–35.

SUTHERLAND, N. S., CARR, A. E. & MACKINTOSH, J. (1962). Visual discrimination of open and closed shapes by rats. I. Training. *Quart. J. Exp. Psychol.* **14**, 129–39.

WILCOXON, F. (1949). *Some Rapid Approximate Statistical Procedures*. New York: American Cyanamid Co.

YOUNG, J. Z. (1960). The visual system of octopuses. 1. Regularities in the retina and optic lobes of octopus in relation to form discrimination. *Nature, Lond.*, **186**, 836–9.

Brit. J. Psychol. (1964), 55, 1, *pp.* 49–57
Printed in Great Britain

THE EFFECT OF VIEWING CONDITIONS AND INSTRUCTIONS ON SHAPE JUDGEMENTS

BY A. A. LANDAUER
Department of Psychology, University of Sydney

A study using different groups of naïve subjects in forty-eight experimental conditions showed that instructions and viewing conditions significantly affected shape judgements. Apparent and 'non-directing' instructions resulted in similar judgements which were as stable as judgements made by subjects who received objective (real shape) or projective (visual angle) instructions. Judgements given under apparent and non-directing instructions tended to follow the visual angle expectation under reduced viewing conditions and under conditions of extreme slants. Under improved viewing conditions and with moderate slants they followed the real shape expectation.

I. INTRODUCTION

Perceptual judgement is determined by a number of factors which include both stimulus features and subject variables such as the 'attitude' induced by the instructions.

Most recent studies dealing with the effect of instructions and viewing conditions have been concerned with size judgements (for a review of some recent work see Epstein, Park & Casey, 1961). Recently, however, Joynson & Newson (1962) and Epstein, Bontrager & Park (1962) have dealt with the effect of instructions on shape judgements. A review of the earlier literature was made by Epstein & Park (1963). Most shape-judgement studies concerned with differences due to varying instructions used only two kinds of instructions. These were either *projective* (analytic, retinal image shape, visual angle shape) or *objective* (real shape, physically equal shape) instructions. The wording of these instructions emphasized that subjects should direct attention to the retinal stimulus, or, in the case of objective instructions, to partly disregard the retinal stimulus.

Thouless (1931) maintained that there are three possible ways in which a subject can judge the shape of a stimulus (distal object): (1) the real shape, that is physical equality; (2) the projected shape, that is, for instance, the exact shape which, held at arm's length, would just obscure the stimulus; and (3) the phenomenal or apparent shape.

Holaday (1933), when investigating the effect of motivational set on size constancy, used four sets of instructions. As well as two sets of projective (analytic) and one set of objective instructions, he also requested his subjects to report the size of the stimulus in terms of the way it is ordinarily seen. Joynson & Newson (1962) gave their subjects what could best be termed 'non-directing' instructions. They did not enlarge on the type of judgement required by their subjects. Joynson & Newson's results show that under non-directing instructions in unrestricted viewing conditions subjects tend to make objective judgements unless the tilt of the stimulus becomes extreme.

The evidence indicates that, provided the stimulus is not viewed under highly reduced conditions, the response of subjects varies as a function of projective or objective instructions. It is not clear, however, whether significant changes in judgements occur with apparent or non-directing instructions.

While it is relatively easy to word instructions which define the operation which a subject has to perform in making a projective or objective shape judgement, the wording of clear instructions for an apparent shape judgement is difficult. Subjects must be induced to judge only the shape as it *seems* or *appears* to them, and not to correct it to the projected or real shape. In non-directing instructions no reference is made to the ways in which judgements can be given.

Over (1960) has shown that size judgements under highly reduced viewing conditions approximate the visual angle expectation when either projective or objective instructions are given. It is reasonable to assume that similar judgements would have resulted had his subjects received apparent, non-directing, or, for that matter, any other kind of instructions which require a size judgement of some sort. It is postulated that under reduced viewing conditions a shape judgement would follow the visual angle expectation, whatever shape judgement instructions subjects receive. As viewing conditions improve the response of subjects will tend to approach the real shape expectation if they receive objective or apparent instructions. Gibson (1950), as well as other experimenters, found that under normal viewing conditions and objective instructions the size of the stimulus can be judged correctly, while under projective instructions a slight regression towards the real size of the stimulus occurs. In general, however, such judgements tend to follow the visual angle expectation. It is reasonable to assume that a similar tendency would be shown in shape judgements.

When a subject is given instructions to report the apparent shape of a stimulus whose orientation varies, his response may differ from the response he would have given had he received projective or objective instructions. Judgements should increasingly approximate the visual angle expectation with increase in the slant of the stimulus and with increasing reduction in viewing conditions. A similar trend should occur when the subject receives non-directing instructions. It may be considered from an operational point of view that apparent and non-directing instructions will lead some subjects to give an objective judgement and others to give a projective judgement. Thus a mean value computed from a sample of subjects would be the confounding of two discrete types of judgement. If this interpretation were correct, the variance of these judgements would be larger than the variance of judgements given under projective or objective instructions.

One of the principal aims of this experiment was to establish whether such differences in the variance occur under different instructions. A further aim was to test the hypothesis that with either apparent or non-directing instructions judgements made under reduced viewing conditions would tend to approximate the projected shape expectation, and judgements made under improved viewing conditions would tend to conform more with the real shape expectation.

II. EXPERIMENT

(i) *Apparatus*

An ellipse with a vertical major axis of 8 in. and a horizontal minor axis of 7 in. served as the stimulus object. It was cut from thin aluminium sheeting and was painted flat mid-grey. The stimulus was attached to an upright black rod and was presented inside a hardboard box 30 in. wide, 37 in. high and 36 in. deep, which was standing on legs 24 in. high. A black curtain served to mask the aperture from view between trials. The slant of the ellipse, which was always in the same direction, could be adjusted by rotating the rod on which it was mounted. Illumination was provided by lamps housed on top of the viewing box and was varied for the different experimental conditions.

The matching shape was a patch of light projected on a milk-glass screen. By turning a small wheel on a shaft extending from the apparatus the subject could alter the ellipticity of the patch of light with respect to its horizontal axis, the vertical axis remaining constant at 4 in. The horizontal width could be varied from 5 in. to almost 0 in. Variations in the width of the matching shape followed a sine function. The reason for using a standard and a matching variable of different linear dimensions was to compel the subject to make shape rather than size matches.

(ii) *Subjects*

A total of 182 males and 106 females acted as unpaid subjects. All were University students and the majority of subjects had no formal training in psychology; none had attended more than an elementary course in psychology. No subject wore glasses during the experiment and none were aware of the purposes of the experiment.

(iii) *Experimental procedure, viewing conditions and instructions*

The experiment was conducted in a dark room. The stimulus was always presented on the left of the seated subject. The distance from the subject's eye to the centre of the stimulus was 100 in., while the distance to the centre of the image on the shape variable was 90 in. Thus a circular stimulus with a diameter of 8 in. in the fronto-parallel presentation would have subtended a solid visual angle of 4·6°, and the solid visual angle of a circular image on the shape variable was 2·6°. The distance between the centre of the stimulus and the centre of the image on the shape variable was 24 in. The centres of the stimulus figures were approximately at eye level.

A. *Viewing conditions*

Three different viewing conditions were used:

(1) *Reduced.* The stimulus was viewed monocularly through an irregularly shaped aperture against a black surround. Head movement was controlled by means of a chin-rest. The stimulus was illuminated by a 125 W. ultra-violet mercury discharge lamp (type Philips 57236 E/70). For this viewing condition the stimulus was covered by a sheet of white paper which was slightly fluorescent. The stimulus was slightly violet in hue and its luminance, measured with an S.E.I. photometer, was 1·30 ft.-lamberts.

(2) *Improved.* The stimulus was viewed binocularly through the same aperture as in condition (1). Head movement was controlled by the chin-rest. The background of the stimulus was a white raised lattice made from ¾ in. wide and ½ in. thick wooden strips. The stimulus thus appeared against a checker-board pattern of ¾ in. raised white squares and 1 in. recessed black squares. A tennis ball was mounted on ¼ in. wooden doweling in the right foreground. A golf ball was fixed to the upper right

part of the background, and at the rear, on the left of the stimulus, a half-opened packet of cigarettes was mounted on wooden doweling. One 25 W. incandescent lamp provided the illumination, with light falling both on the stimulus and the surround. The luminance of the stimulus was 1·70 ft.-lamberts.

(3). *Open.* The stimulus was viewed binocularly through a rectangular aperture 28 in. wide and 22 in. high. Head movement was not controlled. The checker-board pattern formed the background and the remainder of the inside of the box was covered with pink- and white-striped wall-paper. The pink stripes were $\frac{3}{8}$ in. wide and the white stripes were 1 in. wide. In addition to the objects mentioned in viewing condition (2), books, match boxes, pencils, a ruler and a loosely crumpled towel were placed inside the box. A 200 W. incandescent lamp provided the illumination and there were a number of contrasting shadow cues. The luminance of the stimulus was 19·98 ft.-lamberts.

B. *Instructions*

At the beginning of each session one of the four instructions which had been recorded on tape was played to the subject. The demonstration implied in the instructions was given by means of a small model of variously shaped ellipses and a table-tennis bat. Four sets of instructions were recorded for viewing conditions (1) and (2), and four sets for viewing condition (3). Since instructions constitute the main variable in this study, they are reported in full below:

(*Instructions* 1, 2 *and* 3). In the big box straight in front of you, you will see an object. Its surface is flat and it could be shaped like any one of these, but it will be larger. The object may either be perpendicular to your line of sight, or it may be slanted to your line of sight. On the milk-glass screen, on the right of the big box, a patch of light will be thrown and by turning the wheel on the long rod near your right hand the shape of this patch of light can be altered. Your task is to set the patch of light on the milk-glass screen so that its shape:

> (*Instructions* 1)...is the same as the actual shape of the object in the big box. Try to judge the real shape of the object, that is the shape you would obtain if you were to lay the object flat on a piece of paper and draw around the edges with a pencil.
>
> (*Instructions* 2)...is the same as the projected shape of the object in the big box. Try to judge the projected shape, and by projected shape one means the shape of the imaginary patch on the background which is obscured by the object in its present orientation. For instance, if you were to imagine that light was coming from your eye, then the projected shape would be the shadow which the object would cast on the background. The shadow would become narrower, the greater the slant of the object.
>
> (*Instructions* 3)...matches your first impression of the shape of the object in the big box. Try to judge the apparent shape of the object, the way it looks at first sight, without making any allowance for any other information which you might have about it and its position in the big box.

[Keep your head still when making your shape judgement: the head-rest will swivel so that you can face the big box and the screen without lifting your head.]

Don't worry that the light patch is smaller than the object in the big box, the shape can still be perfectly matched. When you are satisfied with your judgement please say the word 'right'. I will then record your judgement and in turn will say 'right', and when I have done so please move the adjustment wheel back so that no light appears on the milk-glass screen. If you are uncertain of the:

(*Instructions* 1)...real shape of the object make the best judgement you can.

(*Instructions* 2)...projected shape of the object make the best judgement you can.

(*Instructions* 3)...apparent shape of the object make the best and quickest judgement you can.

Don't make wild guesses but be as exact as you can.

(*Instructions* 4). In the big box straight in front of you, you will see a flat object. On the milk-glass screen on the right of the big box a patch of light will be thrown and by turning the wheel on the long rod near your right hand the shape of this patch of light can be altered. Please adjust the patch of light on the milk-glass screen so that it looks the same as the object in the big box. When you are satisfied with your judgement please say the word 'right'. I will then record your judgement and in turn will say 'right', and when I have done so please move the adjustment wheel back so that no light appears on the milk-glass screen. Don't make wild guesses but be as exact as you can.

The words in square brackets were included in all four instructions when viewing conditions (1) and (2) were used. They were excluded in all instructions when viewing condition (3) was used.

C. *Slant of stimulus*

The stimulus was presented in either of the following positions: (1) in the fronto-parallel position, 0°, (2) at an angle of 37°, (3) at 53°, and (4) at 66·5°. These angles were selected since they approximate an equal interval step of 0·20 of the cosine of the angle.

Six subjects were used for each of the forty-eight experimental conditions and each subject made five judgements. The angular reading of the subject's judgement on the shape variable was converted into the sine function of the angle and multiplied by ten. This value gave the relative width in inches of the stimulus and was used in the evaluation of the experimental results.

III. RESULTS AND DISCUSSION

(i) *Mean judgements*

The mean judgements for the forty-eight different experimental conditions (based on thirty trials each), as well as the mean judgements for the three first-order interaction effects are presented in Table 1.

An analysis of variance of the data is shown in Table 2. An extension of a model developed by Sutcliffe (1958) was used, which, by unconfounding the intra-subject variation from the inter-subject variation in a double replication design (subjects and trials), gives a very powerful test.

A. Mean judgements, based on 480 trials each, for the three viewing conditions, for all angles of slant and under all instruction conditions were:

(1) 4·878 in. under reduced viewing conditions;

(2) 5·560 in. under improved viewing conditions;

(3) 5·611 in. under open viewing conditions. The difference between means (1) and (2), and means (1) and (3), calculated by Lindquist's (1953) formula 20, were significant at the 0·01 level. Although there was no marked difference in mean judgements when the stimulus was presented in the fronto-parallel position, as an inspection of the relevant means in Table 1 shows, none the less judgements under

Table 1. *Mean judgements in inches*

A. Viewing conditions	B. Instruction conditions	C. Slant conditions				Interaction A × B
		1	2	3	4	
1	1	7·28	5·43	4·06	3·12	**4·972**
	2	6·93	5·26	4·05	2·83	**4·766**
	3	6·75	5·58	4·14	2·79	**4·816**
	4	6·76	5·69	4·43	2·96	**4·959**
Interaction: A × C		**6·928**	**5·489**	**4·170**	**2·925**	
2	1	6·82	6·66	7·12	6·54	**6·785**
	2	6·72	5·54	4·21	2·71	**4·797**
	3	6·66	5·96	5·12	3·12	**5·215**
	4	6·95	5·75	5·02	4·05	**5·443**
Interaction: A × C		**6·788**	**5·979**	**5·368**	**4·105**	
3	1	6·59	6·98	6·80	6·50	**6·719**
	2	6·90	5·48	4·41	2·27	**4·763**
	3	6·72	6·17	5·85	3·45	**5·547**
	4	6·78	6·26	5·50	3·13	**5·417**
Interaction: A × C		**6·748**	**6·220**	**5·638**	**3·840**	
Interaction: B × C						
	1	6·898	6·357	5·993	5·837	
	2	6·849	5·426	4·222	2·604	
	3	6·710	5·902	5·038	3·120	
	4	6·849	5·900	4·981	3·382	

Table 2. *Analysis of variance*

Source	S.S.	D.F.	M.S.	*F*
A. Viewing conditions	160·86	2	80·43	54·31
B. Instructions	365·36	3	121·79	82·23
C. Slant of stimulus	1990·62	3	663·54	448·00
Interaction: A × B	142·99	6	23·83	16·09
Interaction: A × C	113·12	6	18·85	12·73
Interaction: B × C	218·86	9	24·32	16·42
Interaction: A × B × C	141·04	18	7·836	5·29
Error for experimental conditions				
Subjects	355·36	240	1·481	14·39
Error for subjects, trials	118·55	1152	0·103	
Total	3606·76	1439		

Probability of all F-ratios: $P < 0{\cdot}001$.

reduced viewing conditions were over-all significantly smaller than those given under improved or open viewing conditions.

B. Mean judgements, based on 360 trials each, for the four instruction conditions for all angles of slant and under all viewing conditions were:

(1) 6·159 in. under objective instructions;

(2) 4·775 in. under projective instructions;

(3) 5·193 in. under apparent instructions;

(4) 5·273 in. under non-directing instructions. Again, as an inspection of the relevant means in Table 1 shows, there is no marked difference in mean judgements when the ellipse was presented in the fronto-parallel position under the varying viewing and instruction conditions. However, as slant and instructions were experimentally changed, mean judgements varied and an over-all significant difference at the 0·01 level was found between all instruction means, except the difference between means (3) and (4). Thus apparent and non-directing instructions result in similar judgements. This observation was further confirmed by a separate analysis of part of the experimental data. Judgements given under instruction conditions (3) and (4), in all viewing conditions and in slant conditions (2), (3) and (4) were re-analysed. The results of this subsidiary test fully confirmed the findings of the main study: no significant difference was found in judgements made under either apparent or non-directing instructions, and the significant interaction effects involving instructions disappear.

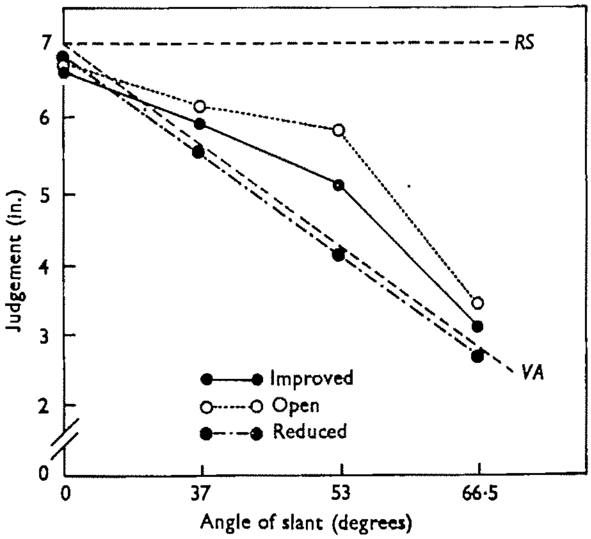


Fig. 1. Mean judgements in inches made under apparent instructions for three conditions of viewing: reduced, improved and open. The plots labelled RS and VA refer to expectation for real shape and visual angle judgements.

C. Mean judgements, based on 360 trials each, for the four angles of slant in all viewing and instruction conditions were: (1) 6·821 in. in the fronto-parallel position; (2) 5·896 in. at an angle of 37°; (3) 5·059 in. at 53°; and (4) 3·623 in. at 66·5°. Although mean judgements did not differ markedly for the varying degrees of slant when given in open and improved viewing conditions under objective instructions (see Table 1), the over-all differences between all means were significant at the 0·01 level.

It may be seen from Table 2 that all interaction effects of the experimental conditions were significant at the 0·001 level. In the statistical model used these interaction effects are independent of the main effects and of each other.

(ii) *Inter-subject variances*

Table 3 presents inter-subject variances other than that part arising from measurement replication. In order to determine whether there was any association between the variability of judgements and instructions given, a non-parametric test was made. The variances of the forty-eight experimental conditions were ranked and these rankings were divided into quarters. A 4 × 4 table containing the cross-classification of instruction conditions and quarters was constructed. The χ^2 for this table, based on 9 degrees of freedom, is 4·67, which shows no departure from chance expectation.

Table 3. *The inter-subject variance for the forty-eight experimental groups*

A. Viewing conditions	B. Instruction conditions	C. Slant conditions 1	2	3	4
1	1	1·544	1·860	0·373	1·329
	2	0·359	5·845	0·613	0·496
	3	0·261	0·993	0·691	0·180
	4	0·988	0·343	0·250	0·216
2	1	0·301	0·955	2·793	8·193
	2	0·141	0·894	1·354	1·971
	3	0·456	0·577	2·056	2·401
	4	0·371	0·041	0·844	12·102
3	1	0·112	2·065	1·850	0·454
	2	0·242	0·350	0·286	0·451
	3	0·978	0·504	5·315	3·372
	4	0·309	0·441	2·231	0·412

This χ^2 value is based on all experimental conditions, including those which showed no marked differences between mean judgements under the different experimental conditions. A similar 4 × 2 table in which the variance of the twelve conditions of fronto-parallel presentation were excluded, and which contained the cross-classification of quarters and either the two operationally or the two non-operationally definable instruction conditions, had a χ^2 of 3·17. With 3 degrees of freedom this again showed no departure from chance expectation.

It is possible that under certain specified viewing conditions there are some real differences between the variances of mean judgements made under the two operationally definable instruction conditions as opposed to the variance of mean judgements made under the two other instruction conditions. There is some indication in Joynson & Newson's (1962) report that this might well be the case. However, as far as this experiment is concerned, such a departure from an over-all insignificant trend in the magnitude of the different variances would have to be specific to only very few experimental conditions. Therefore, the results confirm the hypothesis that apparent and non-directing instructions result in stable judgements which are independent of judgements made when either objective or projective instructions are given. Apparent shape judgements, which are similar to judgements made when non-directing shape-judgement instructions are given, shift from visual angle expectation towards real shape expectation as slant and viewing conditions are varied. However, these judgements do not reach the values obtained when either projective or objective instructions are given.

The results clearly indicate that subjects are able to make meaningful shape judgements if they are given apparent or non-directing instructions. The inter-subject variance of these judgements is not markedly different from the variance of judgements made when operationally definable instructions are given. Thus, in the stipulated viewing conditions, mean judgements obtained under apparent or non-directing instructions seem not to be the result of some subjects giving objective and others projective judgements, but rather a unique type of response, significantly different from responses given under either objective or projective instruction conditions. Although a contrary inference may be drawn from the results obtained by Joynson & Newson (1962), and although the present results may be peculiar to the experimental conditions employed, the view that the interaction of instructions and viewing conditions is an important factor in shape judgements has been sustained here.

This study was conducted while the author held a Commonwealth Post-Graduate Research Scholarship. The assistance of Professor W. M. O'Neil is gratefully acknowledged.

REFERENCES

EPSTEIN, W., BONTRAGER, H. & PARK, J. (1962). The induction of nonveridical slant and the perception of shape. *J. Exp. Psychol.* **63**, 472–9.

EPSTEIN, W. & PARK, J. (1963). Shape constancy: functional relationships and theoretical formulations. *Psychol. Bull.* **60**, 265–88.

EPSTEIN, W., PARK, J. & CASEY, A. (1961). The current status of the size-distance hypothesis. *Psychol. Bull.* **58**, 491–514.

GIBSON, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton Mifflin.

HOLADAY, B. E. (1933). Die Grössenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen. *Arch. ges. Psychol.* **88**, 419–86.

JOYNSON, R. B. & NEWSON, L. J. (1962). The perception of shape as a function of inclination. *Brit. J. Psychol.* **53**, 1–15.

LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin.

OVER, R. (1960). The effect of instructions on size-judgements under reduction-conditions. *Amer. J. Psychol.* **73**, 599–602.

SUTCLIFFE, J. P. (1958). Error of measurement and the sensitivity of a test of significance. *Psychometrika*, **23**, 9–17.

THOULESS, R. H. (1931). Phenomenal regression to the real object. Parts I and II. *Brit. J. Psychol.* **21**, 339–59; **22**, 1–30.

(*Manuscript received* 9 *November* 1962)

Brit. J. Psychol. (1964), 55, 1, *pp.* 59–73
Printed in Great Britain

RESPONSE INSTRUCTIONS AND THE SPEED OF RELATIVE JUDGEMENTS

I. SOME EXPERIMENTS ON BRIGHTNESS DISCRIMINATION*

By R. J. AUDLEY and C. P. WALLIS
University College London

It is hypothesized that the speed with which a relative judgement is made between two stimuli on any given dimension depends upon the relation between the general position of the stimuli on the dimension and the direction of the required judgement along this. Experiments involving the discrimination of light patches of different luminance are reported which support this hypothesis. They show that subjects are quicker at choosing the brighter than the darker of two relatively intense stimuli but quicker at choosing the darker of two relatively dark stimuli. A provisional explanation of the phenomenon in terms of a response competition theory of choice behaviour is offered. Some alternative hypotheses are considered and analogies are made with other psychological phenomena. It is concluded that this is a judgemental phenomenon of some considerable generality.

Shipley, Coffin & Hadsell (1945) showed that when subjects were required to choose which of two colours they preferred they were relatively quick at this task when both colours were high on their scale of preference and relatively slow when the colours were low on this scale. This result could be interpreted as demonstrating that aesthetic judgements are more easily achieved with preferred than with unpreferred material. An alternative explanation involves the possibility of a more general phenomenon of judgement. This would be that *the speed with which a relative judgement is made between two stimuli on any given dimension depends upon the relation between the general position of the stimuli on the dimension and the direction of the required judgement along this*. For example, one would be quicker at choosing the lighter than the heavier of two relatively light weights. In a second paper Shipley, Norris & Roberts (1946) supported this viewpoint by showing that when subjects were required to choose the *least* preferred item they were now relatively quick at this when dealing with colours *low* on their preference scale and relatively slow with those high on this scale.

If this is a phenomenon which is common to all judgements it deserves closer examination. From a theoretical point of view the matter is also of interest because it is a result which might be expected on the basis of a rudimentary theory of choice previously outlined by Audley (1960*a*). However, before dealing with theoretical questions it is desirable to determine whether any phenomenon exists to be discussed. The results obtained by Shipley and his collaborators were not very striking and derived only from studies of aesthetic preference. Therefore a series of experiments was devised to test the hypothesis expressed above in italics. This paper reports some experiments to test the more specific hypothesis that subjects are quicker at choosing the brighter than the darker of two relatively bright stimuli but that the opposite result holds for two relatively dark stimuli. The expected phenomenon is graphically portrayed in Fig. 1.

---

* The research reported in this document has been sponsored by, or in part by, the 'Office of Scientific Research, OAR' through the European Office Aerospace Research, United States Air Force.

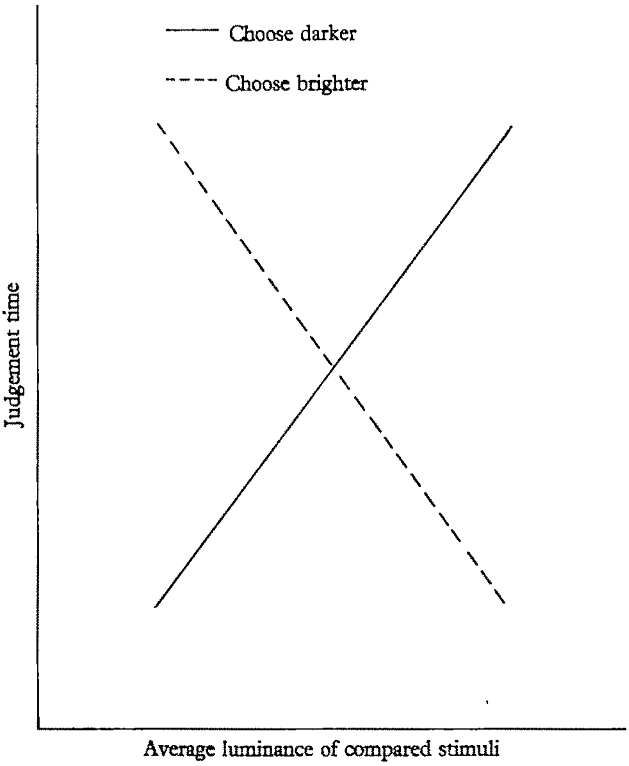


Fig. 1. Predicted interaction between response category and relative stimulus values in a discrimination task.

EXPERIMENT 1

Subjects

The subjects were eight young adults, students at University College London.

Apparatus

A modified Pickford anomaloscope (Pickford, 1957) was used to provide a display consisting of two circular patches of white light each ½ in. in diameter separated horizontally by a gap of 1 in. The seated subject looked through a small circular aperture situated 3 ft. from the display. A shutter was placed across the aperture and an electric chronoscope was started when the shutter opened. The chronoscope was stopped when the subject pressed either of two response keys which were placed at convenient positions for the left and right hands. The experiment was conducted in a darkened room.

Design and procedure

The subjects were required to make relatively difficult brightness discriminations at three levels of display luminance. These were 1·1, 3·0, and 13·8 foot lamberts as measured by an S.E.I. photometer. At each of these levels the two patches of light were set at intensities which preliminary investigations had shown to be difficult to discriminate but not likely to lead to many false judgements. The same objective intensities were used for all subjects.

Each subject made 48 judgements. For half of these he was required to press a key on the side corresponding to the brighter stimulus and for the remainder on the side corresponding to the darker. For each type of relative judgement he was required to make eight judgements at each of the three luminance levels and at every level the correct response was distributed equally between the left and right keys. Half of the subjects made twelve 'darker' judgements followed by 24 'brighter' judgements and finally 12 more 'darker' judgements (*D L L D*). The other half were given the complementary sequence ($L\ D\ D\ L$). This was in order to balance for possible transfer effects from one kind of relative judgement to the other. The time elapsing between the shutter opening to reveal the display and the subject's depression of a key was the dependent variable of the experiment.

Results

The average times for the two kinds of judgement are shown graphically in Fig. 2. On the abscissa of this graph the intermediate level of luminance has been arbitrarily placed halfway between the other two. It can be seen that the expected result portrayed in Fig. 1, which for brevity will from now on be referred to as a *cross-over* effect, was not obtained. Instead, the difference in time between the two kinds of judgement shows a *funnel* effect, being zero at the lowest level of luminance and maximum at the brightest, where the 'brighter' judgements are quicker. A test of the statistical significance of this result was obtained in the following way. The difference between the mean judgement times of each subject for the 'darker' and the 'brighter' responses at the brightest level of luminance was compared with that observed at the darkest level. Seven of the eight subjects showed a greater difference at the brightest level which is a result beyond the 5 % level of significance (sign test, $P = 0{\cdot}035$, 1 tail). The remaining subject showed such a diametrically opposed result that a parametric test fails to reach a satisfactory level of significance.

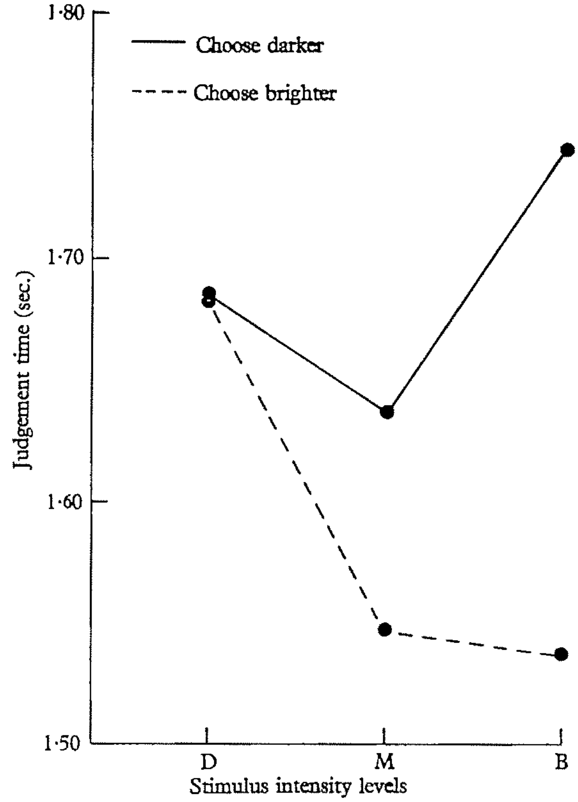


Fig. 2. Expt. 1. Mean judgement times for hard discriminations at three intensity levels brighter than the background.

The variation in the *overall* mean response times as a function of the mean luminance level is presumably dependent upon the relative difficulty of the different discriminations, but there is no reason to suppose that this variation contributed to the observed funnel effect.

The experiment did not provide very conclusive results but it encouraged further experimentation. One simple variation would obviously be to make the discriminations easier. Difficult discriminations had been used in the first experiment because it was felt that this would give more opportunity for the sought-after effects to reveal themselves. It seems more likely that this only served to increase the variance of the data. Therefore in a second experiment subjects were given extremely easy discriminations.

EXPERIMENT 2

Design, etc.

The apparatus and general design remained the same except that easy discriminations were given at only two levels of luminance, the middle level of the first experiment being omitted. A group of eight subjects was again used, but they made twelve discriminations for each of the four possible combinations of type of judgement and luminance level.

Results

The mean decision times for the two kinds of judgement are shown graphically in Fig. 3. The purist might reasonably object to the straight lines drawn between the few points of the graph, but these are only intended to give an easily visualized summary of the results. It can be seen that the funnel effect is again observed. Making the comparison of the differences between judgement times for 'darker' and 'brighter' responses at the two levels of luminance shows all eight subjects relatively quicker with the 'brighter' judgement for the brightest level of luminance. A t-test between the time differences at the two luminance levels gives $t = 3{\cdot}38$ ($P < 0{\cdot}01$, 1 tail).

The differences in the *overall* response times for the two levels of luminance is again probably a function of the difficulty of the two discriminations. The fact that the funnel effect is in the same direction as in the first experiment, although it is *now* the discrimination at the highest level of luminance which takes longer, strongly supports the earlier contention that difficulty is not a principal cause of the effect.

This experiment strongly confirms the result obtained previously and supports the general hypothesis of some interaction between the time required for a particular kind of judgement and the relative position of the stimuli on the dimension being judged. However, a full *cross-over* effect was not found. It could be argued that it is not logically possible to distinguish between the *cross-over* and the *funnel* effects. For example it could be said that any condition serving to affect generally the speed of only one kind of response would shift that particular limb of the graph up or down and turn a *cross-over* into a *funnel* or vice versa. But it seems unlikely that there is anything which could affect the speed of one kind of response which does not fall within the kind of hypothesis which is under examination in this research.

Perhaps there were particular conditions which led to the present result? One strong possibility lies in the background luminance of the displays. Since both experiments were conducted in a darkened room, all stimuli were brighter than the background and consequently it is an obvious possibility, after the event, that the funnel effect is what should have been expected on the basis of the general hypothesis. Therefore a further experiment was conducted in which the display was given a background luminance intermediate between the average of the two brightest and the average of the two darkest stimulus patches.

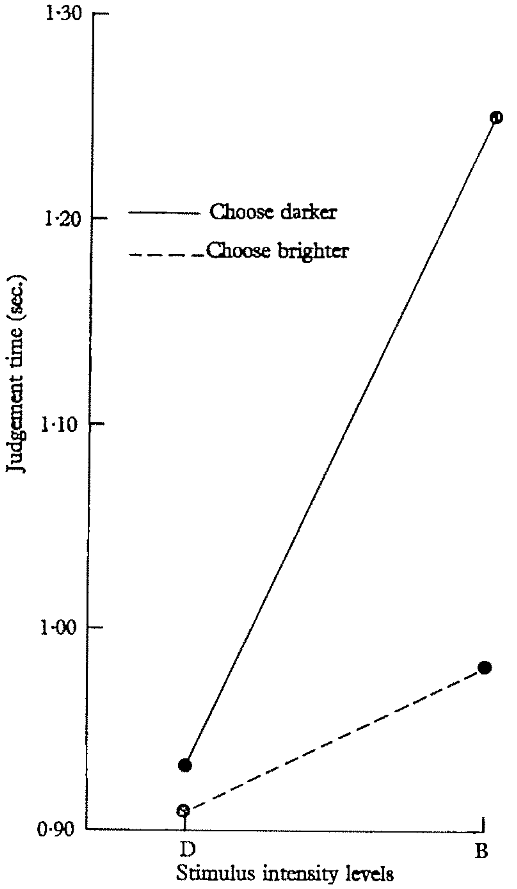


Fig. 3. Expt. 2. Mean judgement times for easy discrimination at two intensity levels brighter than the background.

Experiment 3

Design

The design is exactly the same as that used in Expt. 2. However, a new apparatus was required.

Apparatus

The apparatus was similar in pattern to one used by Hess & Pretori (1894) in a study of brightness contrast. A plan of it is shown in Fig. 4. W_1 and W_2 are wedges of cardboard whose surface can be illuminated by various lamps D_1, D_2, L_3, L_4 etc. Two rectangular apertures, a_1 and a_2, were made in the surface of W_1 so that when S looked through the eyepiece E, he saw a background provided by W_1 and two square stimulus patches provided by the two faces of W_2. The display was equivalent to two 1 in. squares separated by 1½ in. against a circular background viewed at a distance of 2 ft. from the subject. All of the lamps used in the apparatus were 12-volt car bulbs mounted on magnets which could be placed in any required positions on a floor surface of tinplate. By an appropriate system of switches various combinations of lamps could be used to provide the necessary displays. The background was made 2·3 ft.-lamberts and the average levels for the stimuli of the two discrimination tasks were set at 0·6 and 6·0 ft.-lamberts respectively. The stimuli were either both above or both below the background luminance but were always easily discriminable. The experiment was again conducted in a darkened room. Between judgements the subjects rested in the dark. When the display was switched on an electric chronoscope was started and, as in previous experiments, was stopped when a subject indicated his judgement by pressing one of two response keys. The small warm-up period for the electric lamps should not have affected the relative speeds of the various judgements because variation in the luminance of the display was achieved only by changes in the position of the lamps.

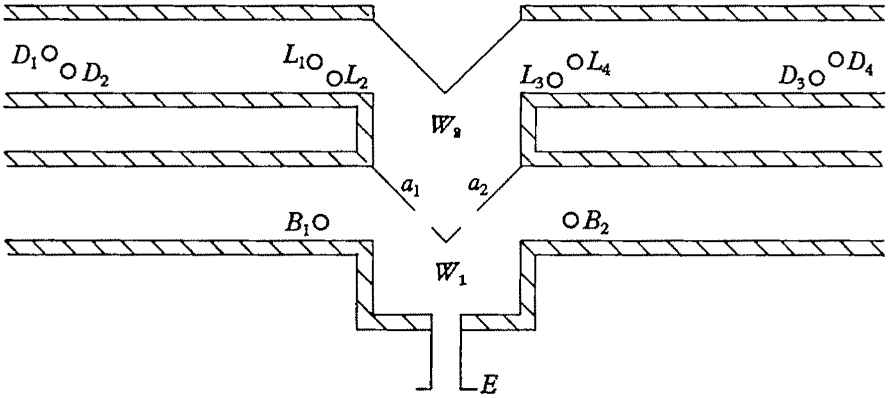


Fig. 4. Ground plan of apparatus used in Expts. 3 and 4.

Results

The mean decision times for the two kinds of judgement are shown graphically in Fig. 5. Again a *funnel* effect is obtained. A test between the time differences at the two luminance levels gave $t = 5\cdot4$ ($P = 0\cdot0005$, 1 tail). It is obvious that the change in the background luminance had no detectable effect upon the result. However, before accepting the *funnel* effect as the most probable result for brightness discriminations it should be noted that the subject rested in darkness between judgements, and this may have been the effective background. In the next experiment the intervening period of darkness was eliminated.

Experiment 4

Design, etc.

The apparatus and design were the same as in Expt. 3. The only difference in the present study is that the background was left illuminated all the time and the stimuli to be discriminated were superposed upon it. The apparatus was not ideal for this purpose, since the apertures where the stimuli appeared had to be unlit, and therefore dark, during the intervals between judgements.

Results

The results of Expt. 4 are shown in Fig. 6. This time a *cross-over* effect has been obtained! Not only is the interaction between the judgement times and the stimulus level significant ($t = 5\cdot5$, $P < 0\cdot0005$, 1 tail), but at each level of luminance the difference in time taken for the two kinds of judgement is also significant (for the relatively dark stimuli $P < 0\cdot05$ and for the brighter ones $P < 0\cdot001$).

It should be noted, however, that the results of this study have been reported and analysed in terms of the overall average times for the different judgements at the two stimulus levels (for Expt. 1, at three levels). This ignores the order in which the various kinds of judgement were made during the course of the experiment. Therefore in Fig. 7 the average times for the two kinds of *stimuli* have been plotted for the four successive blocks of judgements (*L D D L* or *D L L D*) for both Expts. 3 and 4. These graphs show that there is a genuine interaction between the speed of each kind of judgement and the general level of the intensity of the compared stimuli. Clearly, there is a pronounced practice effect; the overall judgement times decrease over the successive blocks of trials. But the times associated with the two stimulus intensity levels change systematically as the response instructions are varied. That the practice effect does not affect the interaction is shown by the fact that in all four graphs the difference in times for the last block of trials is the same as that observed in the first. The two intervening blocks of trials also show no variations in performance. The overall difference between the speed of reaction to the two pairs of stimuli should again be ignored because this reflects differences in discriminability. The *cross-over* effect is therefore a genuine one.

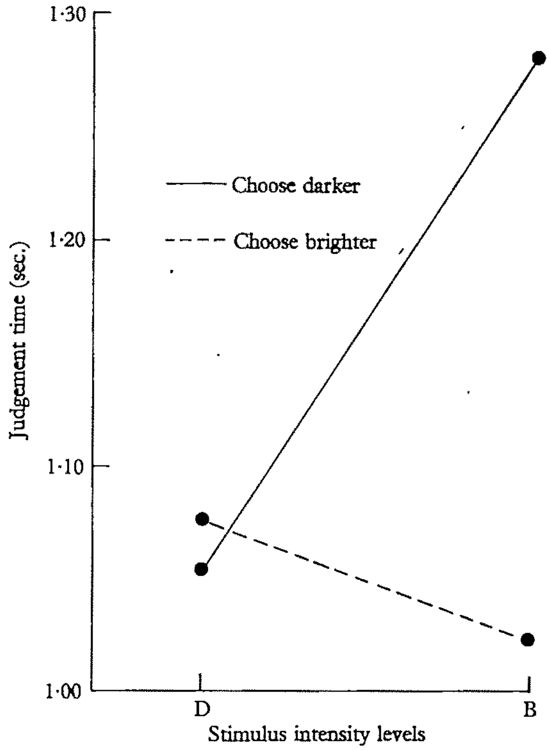


Fig. 5. Expt. 3. Mean judgement times at one intensity level brighter and one darker than the background.

There was a strong implication in the introduction to this paper that it is the nature of the judgement made by the subject which is influencing the time taken to come to a decision. But it may be that, for any given display, the subject tends to glance first at one specific stimulus patch, or in some way to make his judgement with one patch as an initial standard stimulus. For example, he might look at the brighter of two relatively bright stimuli but at the darker of two dark stimuli. This might in some way make it quicker for him to choose the brighter of two bright stimuli, and the darker of two dark ones because fewer scanning movements would be required in these two cases. An hypothesis of this nature might well account for the present results.

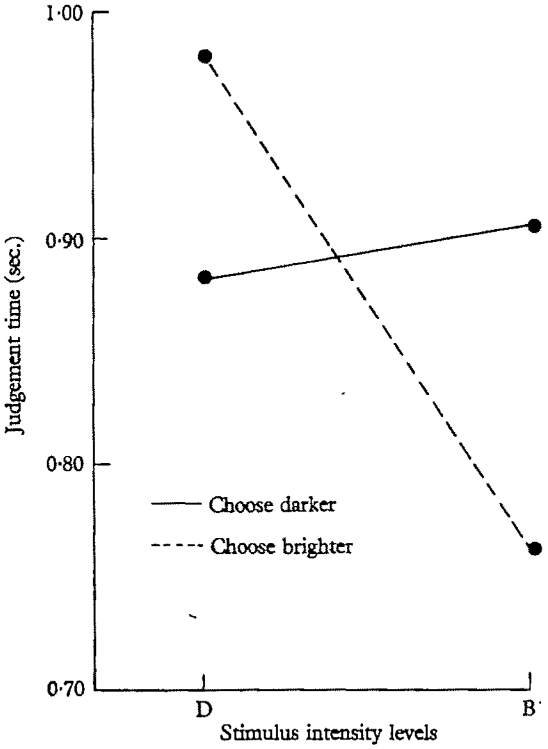


Fig. 6. Expt. 4. Mean judgement time at one intensity level brighter and at one darker than the background with the display illuminated between presentations.

The most obvious test of explanations of this nature is to require subjects to make some comparative judgement concerning the very same stimuli that were used in the previous experiments, but one which is not based upon the relative intensity of the stimuli. In the two control experiments which are next described this was achieved by having a small grey spot present on one or other of the stimulus patches and requiring the subject to indicate in which of the two patches it was located. If any kind of scanning strategy accounts for the previous results it should be equally demonstrated in the discovery of the spot. If, however, it is the relation between the relative judgement and the stimuli being judged which is important there should be no difference between the time taken to locate the spot on the darker and the brighter stimulus, no matter what the general luminance of the pair. Brightness contrast might complicate this conclusion but the luminance of the imposed spot can be made a constant fraction of that of the stimulus patch.

EXPERIMENT 5

Design

Essentially the same design was employed as that used in Expt. 4 except that the task of the subject was to locate the stimulus patch bearing a small dark spot. On one half of the trials the subject was required to press a key on the side of the display possessing the spot (condition *Y*) and on the other half to press a key on the other side (condition *N*). These two response conditions were ordered into blocks of trials in the counterbalanced design common to the experiments reported in this paper. Thus four subjects were exposed to the order *N Y Y N*, with 12 responses

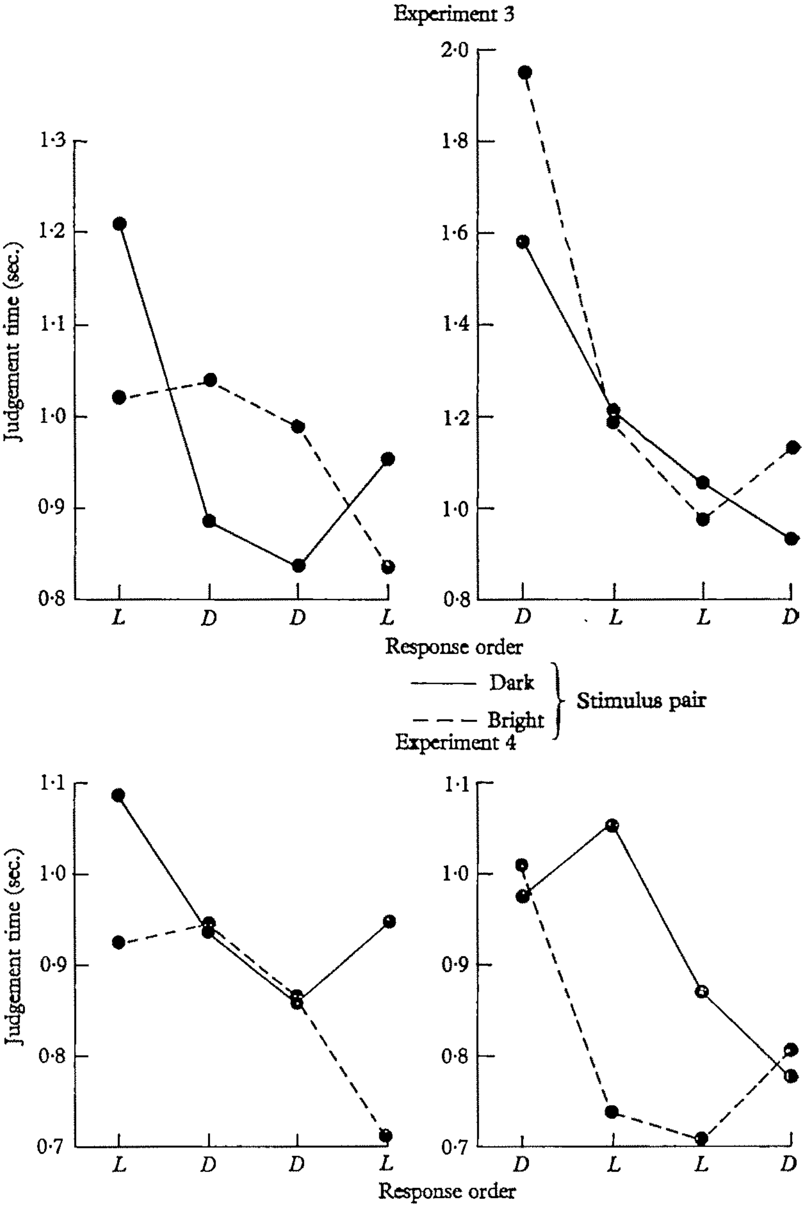


Fig. 7. Mean judgement times at two stimulus intensity levels. Abscissa readings show successive blocks of 'choose lighter' and 'choose darker' responses.

in each block, and another four received the complementary order $Y\ N\ N\ Y$. The ordering of the other stimulus conditions was that used in the previous experiment. The background was kept illuminated between display presentations.

Apparatus

To obtain the spot on one side of the display or the other, the following device was employed. A circular patch of mid grey paper of ½ in. diameter was fixed to each of the faces of wedge W_2 (see Fig. 5) so that they appeared to be in the centres of the two stimulus patches. Two pieces of cardboard of the same optical properties as the wedge W_2, each large enough to cover at least the area represented by each stimulus patch, were attached to the faces of the wedge by means of pivots. By rotation of these cardboard pieces the spot on either or both sides could be eliminated from the display.

Results

Table 1 gives the mean response times for the various displays and response conditions. There is no evidence of any interaction between the difference in the time taken to detect the spot on the brighter and darker patches and the overall luminance of the pair of patches. This is obvious for condition Y, where subjects are on the average 0·03 sec. faster on the brighter side at both levels of stimulus luminance. The relatively slower response to the brighter side at the bright luminance level for condition N does not even approach significance.

Table 1. *Results of Experiment 5. Mean response times at two stimulus intensity levels and for two response categories when the stimulus responded to is the darker or brighter of the stimulus pair*

	Condition Y. Response to side with spot. Relative intensity of pair		Condition N. Response to side without spot. Relative intensity of pair	
Stimulus chosen on:	Dark	Bright	Dark	Bright
Darker side	0·56	0·52	0·55	0·51
Brighter side	0·53	0·49	0·61	0·52

However, although this experiment produced no evidence to refute the interpretation given to the results of the other studies there is a possible flaw in the design. Since the subjects knew that a spot was always present, it was not necessary for them to inspect both stimulus patches before pressing the appropriate response key. If a bare patch was first observed it could be immediately inferred that the spot was on the other side. It is faintly possible that subjects could have responded in this way in the earlier experiments by making a kind of absolute judgement with respect to the brightness of the stimulus patch first seen. But it would seem far more likely that some kind of comparison was involved in the studies requiring a judgement of relative brightness. Therefore in order to force a comparison on the subjects in the spot detection situation it was decided to include some cases in which there was either no spot at all or a spot on both stimulus patches. No decision is then possible unless both patches are inspected.

EXPERIMENT 6

Design, etc.

Only one response instruction was employed. The subjects were required to press the key on the same side as the patch bearing a spot, but only if this was the only spot present. They were to refrain from responding to any other variety of display. For one half of the 48 presentations there was a single spot either on the left or on the right. Of the other 24 presentations one half had no spots at all and one half had spots on both stimulus patches. The values of the stimuli and background were the same as those in the previous experiment, and each possible combination occurred with equal frequency in the four spot arrangements. Eight subjects were again used.

Results

Table 2 gives the mean response times for the various displays. There is clearly no evidence of any interaction between speed of detection and the placement of the spot in the brighter and darker patches. The results of the control experiments are not being used as evidence that subjects do not have particular patterns of searching a visual display. The point is that any such habitual search patterns will not alone account for the results of the first four experiments. It is also necessary to posit some interaction effects between the response required of the subject and the intensity of the stimuli presented for discrimination.

Table 2. *Results of Experiment* 6. *Mean response times in seconds for the detection of one spot*

	Relative intensity of pair	
Response to stimulus on:	Dark	Bright
Darker side	0·72	0·65
Brighter side	0·70	0·66

DISCUSSION

The present studies were carried out because some earlier unpublished investigations of the discrimination of length had suggested that choosing the smaller of two lines might not be simply the psychological opposite of choosing the longer. The two kinds of relative judgement took different times and appeared to be made with different degrees of confidence. This seemed at first to be an anomalous finding because the purpose of these experiments was to explore the relation between the probability of a correct response and the decision times for correct and incorrect choices. Clearly any such purpose could not be unambiguously achieved if the relation between the two response variables changed when the relative judgement required from the subject was changed. Exploration of the literature revealed the study of Shipley *et al.* (1946) as the only relevant paper and so there was an obvious need for further experiments.

On the theoretical side it also seemed possible that some phenomena of the kind reported here could be expected on the basis of a rudimentary theory of choice behaviour (Audley, 1960*a*) to which reference has already been made. It must be admitted that the theoretical argument which follows is not a rigorous derivation from the assumptions of the theory. But it seems to give a first interpretation to the 'facts' and may also lead to further clarifying experiments. To account for a general pattern of relationships existing between a variety of response variables which seemed to hold for many situations (e.g. see Johnson (1955), particularly chap. 10) a simple response competition theory of individual choice behaviour was proposed. This theory may be briefly summarized as follows. Each possible response in any situation is imagined to be continually occurring in covert form at rates of emission particular to the task facing the subject. However, a final, overt choice is not made until the pattern of covert responses is of a specified kind. The pattern considered was an uninterrupted run of covert forms of any one of the responses. On the basis of these simple assumptions it is possible to make a number of predictions about the way in which response variables will vary with one another, either in a simple unchanging situation or in one involving a variety of stimulus displays.

In particular, consider a psychophysical experiment employing the technique known as the method of single stimuli. In this kind of study the subject is presented with a series of stimuli (say, patches of different luminance), one by one, and is required to categorize each of them (say, as 'dark' or 'light'). When the frequency of using one response category is plotted against the stimulus dimension a normal psychophysical ogive is commonly obtained. In order to account for this in terms of the response competition model it would be necessary to suppose that rates of emission of the covert forms of the response 'dark' or 'light' would be related to the stimulus dimension in the manner exhibited in Fig. 8. The linear relations between rate and stimulus intensity should not be taken literally for they are meant to represent any suitable monotonic functions.

In the experiments reported in this paper the method of single stimuli was, of course, not employed. The subjects were required to make a comparison between two stimuli, similar to say S_1 and S_2 in Fig. 8. It is highly probable that with this technique the subject will be able to make use of the presence of both stimuli to sharpen his discrimination beyond the level achieved when each stimulus is presented individually. Nevertheless, the stimuli are usually separated spatially and therefore some succession in comparison is involved whichever psychophysical method is employed. Furthermore, in order to make satisfactory judgements using the method of single stimuli, the subject has to be influenced by the total set of stimuli involved in the experiment, since his judgements are relative to this set. Therefore the hypothetical functions drawn in Fig. 8 do not give response rates which would be appropriate for a series of entirely isolated stimuli but are such as to take into account the complete set of stimuli being judged. In short, context effects operate for both psychophysical methods. Thus it is not too unreasonable, it is being argued, to use the function in Fig. 8 to make predictions about the performance of a subject in the situation requiring a relative judgement about two stimuli.

However, corresponding to each of the two stimuli S_1 and S_2 we have two possible implicit response systems. There is therefore a sense in which we now have a competition between four possible responses which would be expected to be of the kind 'S_1 is brighter', 'S_2 is brighter', 'S_1 is darker', 'S_2 is darker'. If, as in the figure, S_1 and S_2 are both bright, the implicit response rates α_l and β_l associated with the first two statements are much greater than α_d and β_d which are associated with the 'dark' response system. The two 'bright' responses might therefore be expected to dominate any response competition. In fact if one shows a subject a pair of relatively bright stimuli and asks him to say something about the display, he does tell you which is the brighter of the two. On the other hand, he will select the darker of the two dark stimuli for comment. This is just what one would expect on the basis of the above hypothesis (see also Berlyne, 1950).

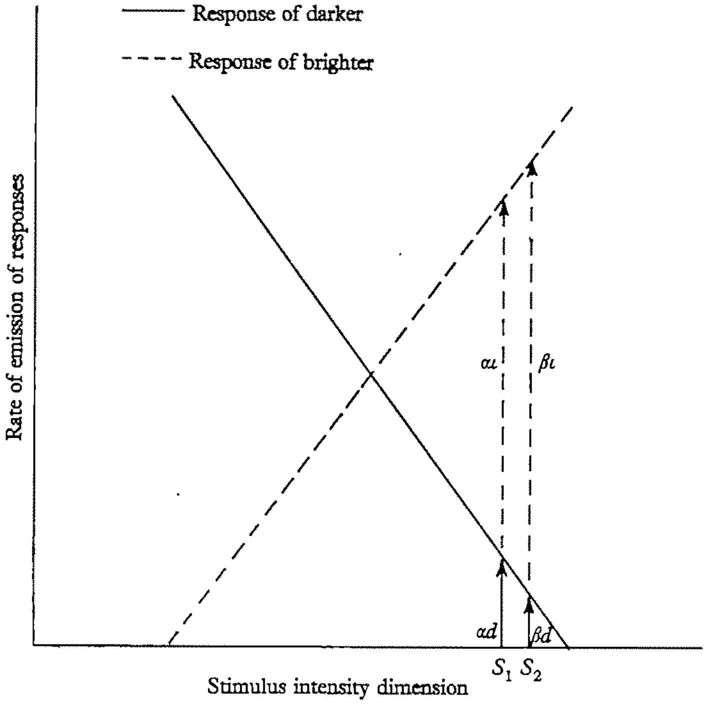


Fig. 8. Rate of generation of covert darker and brighter responses as a function of the relative intensity of the judged stimuli.

When it comes to the time taken to make a judgement under a specific response set, it seems reasonable to suppose that the subject will be using the implicit response system associated with the instruction. Because of the differences in the rates associated with the two systems he will produce the cross-over effect which has been the subject of these experimental studies. It is also possible that the subject always uses the dominant response for a given pair of stimuli and then translates his response to match the experimenter's instruction. This would also predict the results obtained for the extreme stimuli used in the present experiments.

The finding of a tendency to use a particular kind of relative judgement with a given level of stimulus intensity appears to extend the notion of stimulus response compatibility (e.g. Fitts & Seeger, 1953). This concept is used to describe the influences upon time and accuracy produced by the spatial correspondence between the stimulus presented and the response demanded. Two factors are usually considered to be operating. First, the degree of similarity between the spatial arrangement of a response control panel and that of the stimulus display. Secondly, the extent to which the direction of the response movement is congruent with the spatial aspect of the associated stimulus event.

The present studies suggest that the direction of stimuli along a stimulus dimension may be treated in a similar way to spatial direction in the displays of the simple sensori-motor skills used to explore the compatibility concept. Performance is facilitated when the relative judgement required by the subject involves a response which is congruent with the direction of the stimuli along the stimulus dimension.

An analogy may also be drawn with animal discrimination studies. Suppose one trains an animal in a black–white discrimination with the black stimulus positive. If the animal is then presented with two white or near white stimuli it will take a long time to choose between them. One need not, of course, conclude that the animal has in some sense been instructed by the original training to choose the lighter of two stimuli. The discrimination between the two light stimuli can be regarded as involving a conflict between two responses which by prior training, and generalization from this, are either at a very weak level of approach or have even become avoidance tendencies (see for example, Brown, 1942).

So far the question of accuracy of performance has not been raised. In the experiments themselves the relatively easy tasks produced very few errors indeed. On the theoretical side, the analogy with stimulus response compatibility would suggest greater accuracy associated with the faster response and the same conclusion would probably hold for the comparison with the animal discrimination studies. However, the theory behind these experiments illustrated in Fig. 8 would seem to suggest the reverse. The probability of a given judgement of any two stimuli is supposed to be a function of the ratio between the rates with which the covert competing responses are emitted. Suppose the linear relation between these rates and the stimulus dimension, portrayed in Fig. 8, be taken as a first approximation to the true state of affairs. Then for stimuli S_1 and S_2, the ratio between α_d and β_d, the rates for the 'dark' response system, is greater than that between α_l and β_l, representing the 'bright' response system. Hence performance should be less accurate with the favoured response.

Clearly, further experiments are required to explore the basis of the phenomenon we have reported in this paper. But we consider that the phenomenon itself is of some generality. Other studies, not yet published, have shown similar effects in the discrimination of pitch. Other indications are to be found throughout the literature in studies concerned with different problems. For example, Castaneda & Worrell (1961) observed similar (statistically non-significant) trends in a study of latency and response vigour in brightness discrimination.

In conclusion these demonstrations can be taken to emphasize a point which has been stressed elsewhere (Audley, 1960*b*), that much of our experimental exploration of the organization of behaviour is based upon the results of acts of choice whose processes are very imperfectly understood.

REFERENCES

AUDLEY, R. J. (1960*a*). A stochastic model for individual choice behaviour. *Psychol. Rev.* **67**, 1–15.

AUDLEY, R. J. (1960*b*). A stochastic model for the act of choice. *Logic, Methodology and Philosophy of Science.* Proceedings of the 1960 International Congress. Edited by Nagel, Suppes and Tarski. Stanford Univ. Press, 1962.

BERLYNE, D. E. (1950). Stimulus intensity and attention in relation to learning theory. *Quart. J. exp. Psychol.* **2**, 71–5.

BROWN, J. S. (1942). Factors determining conflict reactions in difficult discriminations. *J. exp. Psychol.* **31**, 272–92.

CASTANEDA, A. & WORRELL, L. (1961). Differential relation of latency and response vigour to stimulus similarity in brightness discrimination. *J. exp. Psychol.* **61**, 309–14.

FITTS, P. M. & SEEGER, C. M. (1953). Stimulus-response compatibility: spatial characteristics of stimulus response codes. *J. exp. Psychol.* **46**, 199–210.

HESS, C. & PRETORI, H. (1894). Messenden Untn. über die Gesetzmässigkeit der simultanen Helligkeits-Contrasten. *v. Graefes Arch. Opthal.* **40**, 4, 1–24.

JOHNSON, D. M. (1955). *The Psychology of Thought and Judgment.* New York: Harper.

PICKFORD, R. W. (1957). A practical anomaloscope for testing colour vision and colour blindness. *Brit. J. Physiol. Optics*, **14**, 12–26.

SHIPLEY, W. C., COFFIN, J. I. & HADSELL, K. C. (1945). Reaction time in judgments of colour preferences. *J. exp. Psychol.* **35**, 206–15.

SHIPLEY, W. C., NORRIS, E. D. & ROBERTS, M. L. (1946). The effect of changed polarity of set on decision time of affective judgments. *J. exp. Psychol.* **36**, 237–43.

(*Manuscript received* 13 *February* 1963)

Brit. J. Psychol. (1964), 55, 1, *pp.* 75–84
Printed in Great Britain

ACOUSTIC CONFUSIONS IN IMMEDIATE MEMORY

BY R. CONRAD

Medical Research Council's Applied Psychology Research Unit, Cambridge

Sequences of 6 letters of the alphabet were visually presented for immediate recall to 387 subjects. Errors showed a systematic relationship to original stimuli. This is held to meet a requirement of the decay theory of immediate memory.

The same letter vocabulary was used in a test in which subjects were required to identify the letters spoken against a white noise background. A highly significant correlation was found between letters which confused in the listening test, and letters which confused in recall.

The role of neurological noise in recall is discussed in relation to these results. It is further argued that information theory is inadequate to explain the memory span, since the nature of the stimulus set, which can be defined quantitatively, as well as the information per item, is likely to be a determining factor.

I. INTRODUCTION

One of the accepted hazards of immediate memory experiments, using simple material auditorally presented either from tape recordings or 'live', is the possibility that subjects have not correctly heard the stimuli. One tries to minimize this by speaking clearly, using efficient equipment, and by generally optimizing listening conditions. The amount of error then found which is likely to be perceptual rather than recall will almost certainly be relatively quite small. If one is merely concerned with error quantitatively, it will hardly matter. But if for some reason the nature of specific errors is of concern, then it is possible that perceptual errors may introduce into the results a systematic bias which would be difficult to distinguish from systematic recall errors.

There is at present little evidence available on the likelihood and character of perceptual confusions of letters or digits spoken in the kind of good listening conditions usual in immediate memory experiments. Moser & Fotheringham (1960) present a confusion matrix for spoken digits, but spoken against sufficient white noise as to produce 50 % errors. Curry, Fay & Hutton (1960) have shown that, in good listening conditions, letters of the alphabet spoken at ordinary speech sound levels do indeed yield errors of perception. They show that some letters are more intelligible than others, but offer no confusion matrix. There is therefore some experimental support for the statement made in the opening sentence of this paper. But this does not mean that all systematic errors in immediate recall experiments are perceptual—only that they could be

Now the claims of the decay theory of immediate memory demand the existence of partially decayed memory traces. One would expect that such traces would sometimes yield memories which were not exactly correct, but which were not random with respect to the original stimuli—i.e. not guesses. This becomes then a problem of showing that systematic errors occur after the likelihood has been removed that such errors are perceptual. At this stage it hardly matters in which way a persistently recurring error is related to the stimulus it replaces. Nevertheless, two clues help to suggest the kind of test material that might lead to an adequate demonstration. The first is the occurrence of acoustic confusions referred to earlier; these may not be largely perceptual errors after all. The second is common experience of acoustic confusions in long-term memory (e.g. for names) which Woodworth (1938) noted. For instance, Woodworth recalled azalea as aspasia. It would indeed add point if a mechanism of forgetting common to both long- and short-term memory emerged from this study.

The experiment to be reported therefore describes an attempt to demonstrate the presence, in short-term recall, of acoustic confusions which could not easily be attributed to errors of perception. The necessary procedure calls for a two-part investigation. First, a determination of recall confusions occurring with a set of test items visually presented. Secondly, a determination of the perceptual acoustic confusions that occur when listening to the same test items with no memory involved. The point would be made if there was an association between the two sets of confusions. If memory errors were systematic, but the association with listening errors poor, the data would suggest likely associations.

II. METHOD—RECALL

Material

A vocabulary of 10 letters was used. Two groups of letters were sought the spoken names of which (Ay, Bee, See, etc.) had high within-group acoustic confusability, and low between-group confusability. Since there appears to be no suitable published confusion matrix for the intelligibility of spoken letters, the vocabulary was chosen intuitively and tested subsequently. The following letters were used; B C P T V F M N S X.

A set of 120 6-letter sequences was prepared using the above vocabulary, and subdivided into 6 blocks of 20 sequences. The letter order was basically random but with a number of constraints, viz.; no letter occurred more than once in any sequence: within each block of 20 sequences, each of the 10 letters occurred equally often in each serial position, and within each block, every possible digram (successive letter pair) occurred at least once in each possible serial position.

The blocks of sequences were amalgamated into six different pairs of blocks and a *test comprised the* 40 *sequences* of these amalgamated blocks of 20. The sequences were photographed, 1 letter per frame, on 16 mm. film, each sequence preceded by three 'ready' frames carrying a spot.

The lettering used came from a standard photographic captioning set, but the legibility after photographing was tested. Slides were made of the appropriate letters and were shown to groups of postmen (aged 20–50) under the following conditions. Letters (white on black) were flashed singly in random order onto a greyish screen. The duration of exposure was 1/200th sec. with aperture f. 4·5. A medium neutral density filter was interposed between lens and screen. Subjects wrote down each letter as it occurred, guessing if necessary. The viewing conditions of this test were such that the error rate was about 40%. No systematic confusions between letter shapes was found. It was therefore accepted that in the very much better viewing conditions of the main experiment, consistent errors were unlikely to be due to discrimination difficulties.

Procedure

The sequences were visually presented by means of frame by frame 16 mm. film projection. The projector was controlled by an electronic timer pulsing at a rate of 80 per minute. A new letter thus appeared every 0·75 sec. After the sixth letter of a sequence, the experimenter stopped the projector until subjects were ready for the next sequence.

Subjects were 387 trainee telephonists including males and females and aged between 16–50. (Since the results which follow show no effects due to sex or to age, under or over 40, there will be no further discussion of individual differences.) They were tested in groups of about 10, each group being given one test as defined above. The six available tests were done by approximately equal number of subjects.

Subjects were instructed to watch the screen for the appearance of the 6 letters, and at the end of the sequence to write it down in the correct order on printed answer sheets. They were told to guess rather than leave blank spaces, and the 10-letter vocabulary was written on a blackboard and hence available for reference throughout the proceedings. In this way extra-vocabulary intrusions were almost entirely avoided.

Scoring

Sequences were scored letter by letter. A letter not correct for the particular position in the sequence was scored as wrong, regardless of type of error or any subjective impression on the part of the scorer. Those sequences were extracted for further analysis which had one single letter wrong and no other error. Sequences containing any other kind of error were excluded from subsequent analysis. In this way most of the data returned by the subjects were ignored, and this is why so many subjects were used. But what remained were sufficiently numerous and clearly defined. Sequences retained had only a single *substitution* error. If the data had not been restricted in this way, it is unlikely that anything other than 'noise' would have been added. A 10 × 10 confusion matrix was set up showing which letters were written as responses to which stimuli.

III. METHOD—AUDITORY CONFUSABILITY

Test items

Each of 10 (untrained) speakers recorded 6 practice letters and then the randomized letters of one alphabet in a conversational manner at a rate of 1 letter every 5 sec. The sound level of the speech signal was continuously monitored, and an (on average) equal amount of white noise added. The white noise formed a continuous background on the tape, but was not of course heard by the speakers during recording. The finished tape was divided into two halves each containing five alphabets spoken by 3 male and 2 female voices.

Test procedure

Three hundred Post Office employees (male and female, aged 16–60) listened to one or other of the two halves of tape. They were instructed to listen carefully for the signal (every 5 sec.) and write down on prepared answer sheets which letter they thought it was. They were told to guess rather than to leave blanks. Subjects were tested in groups of 30. The overall error rate was 61 %. A lower rate would have been preferred, but of course if subjects were guessing all the time the rate would have been about 96 %. The errors were set up in a 26 × 26 confusion matrix. (The rank order of the intelligibility of the letters was correlated as between (*a*) the two sets of 5 speakers, (*b*) male and female listeners, and (*c*) listeners under and over 35 years of age (sex kept separate). In every case the correlation was significant beyond the 0·001 level.) Clarke (1957) has amply justified the use of deductions about smaller confusion matrices made from a large one. Accordingly the relative acoustic confusability of the 10 test letters was derivable from these data.

IV. RESULTS

The resulting confusion matrix of the auditory confusability test is shown in Table 1. It shows the frequency with which a particular response was made to a particular stimulus letter when the response was incorrect. For general reference, the complete error matrix for the whole alphabet is given as an Appendix.

Inspection confirms that the selection of these ten letters was suitable in that confusions are relatively high within the two groups BCPTV and FMNSX and relatively low between them. It should be remembered that the choice of these 10 letters is quite arbitrary. It is only necessary that there should be some significant confusions amongst them—and this is clearly so. The critical question is whether the letters which confuse in the auditory confusability test, also confuse in short term memory. Table 2 gives the confusion matrix resulting from the memory test for those cases where only a single letter in the sequence was incorrect.

The first point to note is that errors are not random. Inspection indicates that when an error of recall is made, the substituted letter is likely to be one which *sounds* like the correct letter. A 2 × 2 chi-squared test using the sum of the values in the four quarters of Table 2 shows an association of this kind which is significant beyond the 0·001 level. Another way of looking at the data is to consider each stimulus letter and examine the distribution of recall errors amongst the remaining 9 letters. Chi-squared tests indicate that in no case can this distribution be considered random. In other words, when a letter cannot be remembered, what the subject records is not determined by chance. Certain systematic substitutions are made, the nature of which was examined in relation to the listening data of Table 1.

Table 1. *Listening confusions*

	Stimulus letter									
Response letter	B	C	P	T	V	F	M	N	S	X
B	.	171	75	84	168	2	11	10	2	2
C	32	.	35	42	20	4	4	5	2	5
P	162	350	.	505	91	11	31	23	5	5
T	143	232	281	.	50	14	12	11	8	5
V	122	61	34	22	.	1	8	11	1	0
F	6	4	2	4	3	.	13	8	336	238
M	10	14	2	3	4	22	.	334	21	9
N	13	21	6	9	20	32	512	.	38	14
S	2	18	2	7	3	488	23	11	.	391
X	1	6	2	2	1	245	2	1	184	.

Table 2. *Recall confusions*

	Stimulus letter									
Response letter	B	C	P	T	V	F	M	N	S	X
B	.	18	62	5	83	12	9	3	2	0
C	13	.	27	18	55	15	3	12	35	7
P	102	18	.	24	40	15	8	8	7	7
T	30	46	79	.	38	18	14	14	8	10
V	56	32	30	14	.	21	15	11	11	5
F	6	8	14	5	31	.	12	13	131	16
M	12	6	8	5	20	16	.	146	15	5
N	11	7	5	1	19	28	167	.	24	5
S	7	21	11	2	9	37	4	12	.	16
X	3	7	2	2	11	30	10	11	59	.

The association between the two complete sets of confusions was tested by calculating Spearman's coefficient of rank correlation. Considering all 90 cells, a value of +0·64 was obtained. This yields a *t* value of 7·68 which is significant at beyond the 0·0001 level. This establishes beyond any doubt that even with visual presentation of material to be memorized, when recall errors occurred they were similar in nature to hearing errors when the signal was partially masked by white noise.

V. DISCUSSION

Acoustic confusions in recall

It was pointed out in the Introduction that a decaying memory trace ought to be revealed by the presence of errors which bore a systematic relationship to the original stimulus. A conventional memory span test samples the decay curve over the critical period of time when decay is rapid. For a given stimulus, the decaying trace might be expected to yield a different recall response at different points in the decay time-curve. But the decay curve of a given stimulus, always examined at the same point, should lead to consistent erroneous response. Table 2 clearly shows the presence of errors systematic in a manner which would support this analysis.

The Introduction further suggests that errors would be systematically related to stimuli in one particular way; there would be acoustic resemblance. The rank order correlation between Tables 1 and 2 indicates the remarkable similarity between memory errors and listening errors. The confusions reported when listening to noise-masked speech provide a highly reliable prediction of the course of forgetting.

There is of course no mystery in the fact that visually presented memory material should yield acoustically related errors. It merely indicates that the majority of subjects verbalize the stimuli, rather than attempting to store them in visual form. What is interesting is not that the memory distortions correspond to the sound of the original stimuli verbalized, but that they correspond so closely to the distortions occurring when speech sounds are partially masked by white noise. This point will be returned to later.

Indeed the true amount of association must be even higher than that shown. The 'listening' evidence contains a certain amount of random error—both in itself, when subjects have been unable to decide between any of the alternatives, and in its relationship to the 'memory' data. Ideally the association should be tested for each subject individually with his own speech used for the listening task. The small sample of voices used in the listening task would be only an approximate match to any individual voice. These factors would tend to lower the correlation between memory errors and acoustically masked speech.

These results clearly support the functional description of forgetting outlined by Brown (1959) in terms of information theory. For instance, Brown describes decay of a trace as 'a fall in the signal-to-noise ratio'. He vividly illustrates this by an analogy of a letter chalked on a blackboard and increasingly smudged. 'Some smudging...will leave it legible: it starts with internal redundancy so that some "decay" does not matter. Further smudging, however, will make it difficult to read....'

What an informational model inevitably misses, though, is that during the course of 'smudging', a systematic change in the signal-to-noise ratio will yield systematic output changes. The logic of information theory is such that unless specifically stated, alternative responses are equi-probable. A weakness of an informational model of forgetting is that in the Brown-type analogy, a letter is legible in spite of some smudging. But when smudging has increased to a point where the letter is no longer correctly legible, all the letters in the stimulus set are equally likely to be reported.

Acoustic confusions and memory span

This logic has led to the extremely important prediction that memory span is a function of the information per item in the stimulus set (Pollack, 1953; Miller, 1956). The present data suggest an alternative prediction.

The facts seem to be that when an item cannot be recalled, the probabilities of other items in the set being given as response are constant and unequal. Systematically, certain items are more likely to be given than others. In the material being discussed, these probabilities depend on the acoustic similarity between the various items. From this one could argue that the more chance there is of acoustic confusion within the stimulus set, the poorer will recall be. It would follow that the memory span would be a function of the acoustic similarity of the members of the set. The span might depend not on the number of alternative items in the set, but on the number of acoustically similar items in the set. Increasing the size of set might or might not increase the number of acoustically similar items in it.

One can imagine four vocabularies, viz: (1) 2-letter high confusability; (2) 2-letter low confusability; (3) 8-letter high confusability; (4) 8-letter low confusability. On an informational basis, one would expect (1) and (2) to yield the same memory span. Also (3) and (4). One would further expect the span for (1) and (2) to be longer than for (3) and (4). But on the basis of the argument above, the size of the spans, taken to be determined by the acoustic similarity in the sets, could well be in the order: (2) (4) (1) (3).

Experimental evidence on the question of vocabulary size and span is conflicting. Miller (1956) quotes unpublished work by Hayes, supporting Pollack's (1953) earlier study, indicating that span is more or less independent of vocabulary size. Hayes used digits, letters and words; Pollack used digits and letters. In neither case are the actual items given which were used for each size of set. Crossman (1960) using a number of ingenious but unusual vocabularies—e.g. black/white, £. s. d., N., S., E., W., etc.—showed the relationship predicted by information theory. Conrad & Hille (1957) using 2-, 4-, and 8-digit vocabularies found the opposite to Crossman, that the larger the vocabulary the better the span.

Without knowing the exact sets used by Hayes and by Pollack, their results cannot be discussed in terms of confusability. Crossman's could if the acoustic relationships were estimated and if other things were equal between the vocabularies he used. In the case of Conrad & Hille, it is notable that the 2-digit set used the digits 2 and 3, the 4-digit: 2, 3, 4, 5, and the 8-digit: 2, 3, 4, 5, 6, 7, 8, 9. Moser & Fotheringham (1960) have shown that the two digits most likely to confuse acoustically are in fact 2 and 3. Additional digits are likely to make the set in entirety less acoustically confusing, and would on the present argument lead to longer spans. It may be that Miller's colourful statement about the magical number 7, is, for the memory case, right but for the wrong reason. Whether seven items minus or plus two can be recalled may depend not on vocabulary size, but on the nature of the items used in each stimulus set.

Present results lead to the general prediction that memory span is a function of the number of acoustic confusions expected from the vocabulary items used. The expected value depends on the probability with which any item confuses with any other. Strictly, of course, the word confusion here refers to recall confusion whether acoustic or any other. In this case, prediction of the span for a given vocabulary is not possible until after the event when the confusion probabilities are known. But because of the special relationship between the data in Tables 1 and 2, and because therefore Table 1 gives approximate estimates of the probability of recall confusions for alpha vocabularies, the appropriate prediction could in principle be made and the model specifically tested. The next section will discuss this unique relationship between acoustic confusions in listening and confusions in recall.

The role of neurological noise in memory

Brown (1959), more than anyone, has developed the concept of noise as a determinant of forgetting. He equates decay of a memory trace with 'decrease in the signal-to-noise ratio'. In his use of the term noise, Brown is thinking of an information theory model where the noise concept is necessary to account for the discrepancy between input and output. Brown specifically is not thinking of noise as having any physiological correlate. In the information theory sense, the presence of noise will ensure degradation of the signal, but the way in which Brown's functional model is formulated precludes any specification of the degradation. The present experiment has demonstrated a highly specific and consistent distortion of signal. One must therefore suppose the presence of a specific type of noise which is random with respect to the specific nature of the memory trace. Neurological noise might very well meet the requirement.

Hebb (1961) writes: 'Any random activity in these excess neurones (the ones not needed for the task being learned) is "noise", which must tend to interfere with the learning.' In the same Symposium, Eccles (1961) says that 'if there is electrical interaction...everywhere in the nervous system, then everything should be electrically interacted with everything else. I think this is only electrical background noise...I would say that there is electrical interaction, but it is just a noise, a nuisance.'

Two postulates emerge from this: (1) a pattern of neural activity distorted by neural noise; (2) because of the relationship between Table 1 and Table 2, a form of neural storage which preserves directly definable aspects of the physical stimulus. Memory errors are thus seen to be related to the signal–noise ratio prevailing at the time of recall. That the ratio will change is highly likely. But the question whether this is through decrease in signal strength or increase in noise suggests some speculations which may be worth following.

The extremely rapid rate of forgetting in the absence of rehearsal is now accepted. Fairly quickly (in a matter of seconds) the pattern of specific neural activity will dissipate until it is no longer discriminable from the continuous random neural activity which constitutes neurological noise. The memory trace is irrecoverably lost. But many memories have a more permanent neural basis, an inescapable conclusion from the fact that some distant events can be recalled. What is perhaps puzzling is not only the fact of long-term recall, but especially the fact that it is erratic. Temporary failure to recall, for instance, a well-remembered name is common enough. It is not in this context very easy to understand why the signal—a pattern of neural activity—should erratically wax and wane in strength. But it is highly probable that the level of neurological noise changes continuously. If long-term storage followed the basic pattern apparent for short-term memory, an acoustically coded memory trace, such as a name, would suffer the same type of distortion in noise. In attempting to recall the name three possible states would be expected according to the prevailing level of neurological noise. Complete failure to recall (excessively low signal–noise ratio), complete recall (adequate signal–noise ratio), imperfect recall with acoustic relationship to the original stimulus. This third state is critical and is of course everyone's experience.

There are two features of this particular type of recall relevant to the postulate of neurological noise as a factor in forgetting. The fact that first-shot recall is often acoustically similar to the original, but that one is confident that it is wrong; the fact that correct recall, if it is reached, is accompanied by complete certainty that it is right. It seems much easier to imagine a permanently established trace subject to masking by a fluctuating level of neurological noise, than to imagine a process of reconstruction from a trace merely of inadequate signal strength. Any agent which could reduce the level of neurological noise (one assumes that in the long term case it is no longer possible to raise the signal level) would facilitate recall. Russell (1959) refers to psychogenic amnesias which can be recovered by persuasion under the influence of drugs or hypnosis. It would not seem too fanciful to suppose that these and other agents affect neurological noise level. This is perhaps a too simple-minded view. But there is no doubt about the facts of remembering cited, the existence of neurological noise, the peculiar nature of recall errors, and their special relationship with certain types of perceptual error.

VI. CONCLUSION

In so far as evidence claimed to support the decay theory of immediate memory has been adduced, it has always hitherto been quantitative. An imposed delay, without rehearsal, between presentation and recall of material has led to overall more forgetting. In the present experiment, the unit of quantity has been so shifted that one can more conveniently talk of qualitative performance change. The quality, or the nature of the recall response to single-letter stimuli has been shown to change in a consistent manner before the decay point is reached when response is random with respect to the stimulus. Decay theory demands a state of partial decay as well as complete decay. This state has been demonstrated.

That the memory trace, in this kind of situation, has an acoustic or verbal basis, is at least convenient. It permits specific prediction to be made about qualitative changes in short-term forgetting when the nature of the stimulus material is manipulated. Nor is the general relationship shown here between masked listening and forgetting necessarily confined to the hearing sensory mode. In principle, non-verbal material could be examined in the same way. One would expect a visually perceived pattern which could not be named, to be recalled with the same changes that would occur if the pattern were seen with visual-noise masking.

There is clearly a similarity between an information theory view of short-term forgetting most comprehensively described by Brown (1959), and the present outline. But the resemblance is often in the words used rather than in the concepts. In this paper, decay refers to a changed pattern of neural activity in a physical sense; in information theory, decay is a decrease in the signal-to-noise ratio. Of course the latter is a more general statement of the former, but carrying the implication that either signal or noise could change. In both cases though, the term noise refers to a disturbance random in relation to the signal. In the present case, it is necessary to describe this in neurological terms. Again the term redundancy has something in common with the confusability concept, but it lacks the specificity. Informationally, noise and redundancy can only be inferred from an input–output discrepancy. In the present case, noise can (in principle) be measured, confusability be controlled, and the input–output discrepancy predicted. In fact it may be the vagueness of the redundancy concept when applied to memory which has led to the ill-supported prediction that span would be a function of information per item. On the other hand, it is quite obvious that the concept of neurological noise, as it has been used here, is necessarily post-Shannon.

Finally, again, the value of closely examining the nature of errors in human behaviour has been exemplified; especially when the errors can be objectively defined, counted, and treated statistically. An error indicates (the only indication) an imperfectly behaving system. Whilst measuring the amount of error in defined situations is clearly invaluable and often the best course available, the forms of malfunction may provide a short cut to an understanding of structure.

The author is deeply indebted to the Post Office and to the officers of the Union of Post Office Workers for their assistance in providing subjects for this study. The Post Office also provided accommodation. All of the testing and scoring of the recall test was carried out by Miss D. J. A. Longman (G.P.O.) and Mrs A. J. Hull. The listening test was carried out and scored by the G.P.O.'s Central Organisation and Methods Branch.

REFERENCES

Brown, J. (1959). Information, redundancy and decay of the memory trace. In *The Mechanisation of Thought Processes. Nat. Phys. Lab. Symp.* no. 10, H.M.S.O.

Clarke, F. R. (1957). Constant ratio rule for confusion matrices in speech communication. *J. Acoust. Soc. Amer.* **29**, 715–20.

Conrad, R. & Hille, B. A. (1957). Memory for long telephone numbers. *P.O. Telecomm. J.* **10**, 37–9.

Crossman, E. R. F. W. (1960). Information and serial order in human immediate memory. In *Information Theory*. London: Butterworth.

Curry, E. T., Fay, T. H. & Hutton, C. L. (1960). Experimental study of the relative intelligibility of alphabet letters. *J. Acoust. Soc. Amer.* **32**, 1151–7.

Eccles, J. C. (1961). The effects of use and disuse on synaptic function. In *Brain Mechanisms and Learning*. Oxford: Blackwell.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In *Brain Mechanisms and Learning*. Oxford: Blackwell.

Miller, G. A. (1956). The magical number seven plus or minus two; some limits of our capacity for processing information. *Psychol. Rev.* **63**, 81–97.

Moser, H. M. & Fotheringham, W. C. (1960). Number telling. *Tech. Rep.*, no. 58, *RF Project* 1080. Ohio State University Research Foundation.

Pollack, I. (1953). Assimilation of sequentially encoded information. *Amer. J. Psychol.* **66**, 421–35.

Russell, W. R. (1959). *Brain. Memory Learning.* Oxford: Clarendon Press.

Woodworth, R. S. (1938). *Experimental Psychology*. New York: Henry Holt.

(*Manuscript received* 3 *January* 1963)

APPENDIX. LISTENING ERRORS, FULL ALPHABET

Stimulus letter

Response letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	.	25	9	24	18	9	35	77	19	151	200	134	45	126	478	9	8	23	22	17	28	22	79	8	21	22	
B	2	.	171	197	107	2	142	15	13	30	11	3	11	10	4	75	94	—	2	84	33	168	4	2	3	25	
C	2	32	.	19	24	4	19	7	3	10	4	—	4	5	—	35	30	—	2	42	4	20	1	5	5	3	
D	4	125	71	.	166	1	104	4	12	21	6	3	11	7	—	34	47	2	—	43	11	105	4	2	6	11	
E	8	271	195	252	.	11	208	31	100	9	16	7	59	77	9	111	66	6	6	163	65	119	11	19	5	6	
F	10	6	4	4	4	.	8	55	2	12	30	25	13	8	6	2	–	4	336	4	4	3	4	238	5	81	
G	4	40	22	47	34	—	.	2	9	8	5	—	3	3	1	18	39	—	3	12	33	37	2	2	1	1	
H	27	3	15	13	22	26	3	.	1	15	28	7	9	15	34	6	16	4	43	12	2	5	14	28	1	32	
I	1	2	5	2	14	5	4	3	.	14	6	51	52	30	23	1	5	188	21	6	5	3	5	4	78	5	
J	10	5	6	3	1	1	12	8	—	.	158	9	4	2	7	8	7	1	2	3	8	10	8	3	20	35	
K	46	2	7	6	—	7	—	15	7	64	.	16	9	17	21	6	12	7	7	6	1	5	10	3	6	28	
L	69	6	2	3	4	16	6	9	35	65	42	.	32	25	73	4	6	57	38	1	2	3	32	4	61	41	
M	106	10	14	5	13	22	22	9	31	42	73	83	.	334	24	2	7	22	21	3	6	4	17	9	30	41	
N	234	13	21	20	35	32	31	40	47	104	127	51	512	.	59	6	6	40	38	9	12	20	28	14	47	55	
O	116	10	5	10	18	22	26	16	31	80	80	122	15	30	.	9	8	41	26	36	15	20	76	6	15	50	
P	7	162	350	201	172	11	167	21	12	18	39	2	31	23	4	.	231	—	5	505	16	91	2	5	3	31	
Q	3	43	90	59	52	—	103	16	8	12	8	1	3	6	1	215	.	1	1	90	41	26	5	2	3	9	
R	13	3	3	3	13	14	10	4	104	17	20	136	43	12	15	1	9	.	15	4	2	5	7	9	162	28	
S	20	2	18	9	17	488	4	151	3	38	34	22	23	11	18	2	9	5	.	7	3	3	5	391	10	111	
T	3	143	232	158	124	14	116	20	9	11	13	1	12	11	4	281	175	6	8	.	11	50	8	5	2	10	
U	10	106	74	90	125	2	189	26	50	23	24	1	22	23	11	79	112	3	5	29	.	347	13	13	4	8	
V	4	122	61	147	55	1	135	13	6	15	18	—	8	11	—	34	35	1	1	22	79	.	2	—	6	15	
W	4	31	5	23	7	9	4	4	17	5	11	6	8	11	1	6	3	2	2	2	17	19	.	1	14	91	
X	1	1	6	3	6	245	1	163	1	5	12	9	2	1	1	2	1	2	184	2	2	1	2	.	2	56	
Y	7	5	7	1	5	7	11	3	86	98	25	62	39	11	7	1	2	32	5	7	4	5	2	4	.	22	
Z	4	1	5	3	4	40	4	21	2	115	52	6	7	3	3	3	5	2	9	—	4	6	6	22	21	.	
	715	1169	1398	1302	1040	989	1359	733	608	982	1042	757	977	812	804	950	933	449	802	1109	408	1097	347	799	531	817	22,929

Total no. letters presented = 37,440.

Brit. J. Psychol. (1964), 55, 1, *pp.* 85–88
Printed in Great Britain

MIMICRY IN MYNAS (*GRACULA RELIGIOSA*): A TEST OF MOWRER'S THEORY

BY BRIAN M. FOSS
Birkbeck College, University of London

Mowrer considers that talking birds will imitate only those sounds which have been associated with primary or secondary reinforcement. His theory was tested by playing whistles to two groups of myna birds, in the presence of reinforcers and in their absence. The birds mimicked the whistles equally well under each condition.

I. INTRODUCTION

Mowrer (1950, chap. 24) states that sounds must be associated with reinforcement for talking birds to imitate them. In his most recent formulation (Mowrer, 1960) the reasoning is as follows: if any stimulus, for instance the sight of a human being, is repeatedly associated with a primary reinforcer (e.g. food) then the appearance of the human will give rise to 'hope', which in turn is reinforcing; if the human repeatedly utters a given sound, that sound will also produce hope; if now the bird, in the course of its babbling, makes noises which approximate to those produced by the human, these will produce hope, and the production of the noises will be reinforced—the more so the more the noises approach the human version.

Skinner (1957, p. 64) appears to accept this view, although he is not primarily concerned with factors affecting the first appearance of any response. He is concerned with the control of the response once it has appeared, and this has been shown to be possible with the utterances of mynas (Grosslight, Harrison & Weiser, 1962).

Mowrer believes that imitation learning by talking birds can be taken as a model for other kinds of imitation. A child will tend to imitate noises of adults, especially those with whom he is identified, since they will have the greater reinforcing power. Obviously the exact mechanism of learning will be more complicated in cases other than learning to produce sounds. For sounds, the child can match his own utterances against those produced by the person being imitated, but it is not clear how this matching can happen when for example a movement is being imitated.

Some doubt is felt about Mowrer's theory for two reasons: (i) talking birds often pick up noises which are not consistently associated with reinforcement (e.g. the noise of a dripping tap); (ii) bird fanciers often recommend training birds to talk with the cage covered, and also by means of gramophone records (see, for example, *Pet Myna*, 1957), situations which seem to minimize the chances of hope, in Mowrer's sense.

The present experiment attempts a rough test of the theory that learning to utter a sound depends on the sound's being associated with reinforcement.

II. METHOD

Subjects

The subjects were two groups of Indian Hill Mynas (*Gracula Religiosa Intermedia*) bought as 'gapers' and aged about 8 months at the start of the experiment. Group A consisted of two birds and group B of four. This difference in numbers was incidental to the experiment. Since these birds learn from each other, in a sense the experiment was done on two subjects. The two groups were treated slightly differently: group A were kept in a cage in the experimenter's room; group B were all in one cage in a room adjoining another which could be reached by the experimenter without the birds seeing or hearing him.

Procedure

The stimuli used were an ascending whistle (X), going from 500 c.p.s. to 2000 c.p.s., in the space of 2 sec., and the same whistle reversed, and therefore descending (Y). Each whistle was recorded on a loop of tape which repeated the whistle every 6½ sec. This kind of whistle is not heard in the repertoire of an unsophisticated bird and its timbre was also unusual. This choice was made to help in identifying the whistles when the birds finally produced them, and also to minimize transfer. A whistling ban was imposed in the neighbourhood of the cages.

The birds were trained as shown in Table 1.

Table 1. *Training schedule*

	Whistles	
Birds	X	Y
Group A	Without reinforcement	With reinforcement
Group B	With reinforcement	Without reinforcement

This design did not allow for controlling which whistle was heard first for each group. To give Mowrer's theory priority, each group heard first the whistle associated with reinforcement, the number of repetitions of each whistle being gradually increased from 10 to 30 per session. The whistles associated with reinforcement were always played at midday, the others at dusk (Table 2).

Table 2. *Numbers of repetitions of stimuli per session*

	Midday (reinforcement)	Dusk (no reinforcement)
Day 1	10	10
Day 2	20	20
Day 3 onwards	30	30

At the midday sessions food was prepared in full view of the birds, and put in the cages (always by the same experimenter) while the appropriate whistle was being played. At the dusk sessions the groups were treated differently. Group A lived in a room where they saw many people and heard a good deal of conversation. Their cage was covered at least 15 min. before the tape was played and remained covered for at least 30 min. afterwards. During this period they heard no human vocal noises. For group B, living in an isolated room, the tape was switched on in an adjoining room, so that the birds neither heard nor saw a human during the session.

Recording results

After 5 weeks (less 2 week-ends) of this daily procedure, the procedure was stopped and recordings were made of each group, with and without the experimenter in the room. For group A a total of 12 hr. was sampled, for group B 9 hr., both spread over 3 days. The recordings were made at various times between 9.30 a.m. and dusk.

III. RESULTS

The birds tended to incorporate the test whistles into other whistling and babbling, and later (after the experiment was over) tended to construct variations on the original themes. However it was possible to count those whistles which were produced in isolation from other utterances, and which clearly had the correct timbre and the correct ascending or descending character. (The birds were successful in reproducing the whistles over only about half of the range of the test stimuli.) The results are shown in Table 3.

It is clear that both groups had learned both whistles. There was no tendency for one whistle rather than the other to be produced at any particular time of day. Group A produced whistles equally often while the experimenter was in and out of the room; group B never produced the whistles while the experimenter was near during the period under consideration. They did so for the first time several days after the experiment was finished.

Table 3. *Numbers of whistles reproduced over the times sampled*

	Whistle associated with reinforcement	Whistle not associated with reinforcement
Group A	26	21
Group B	11	13

Other findings

It was not until recordings were made in group B's room in the absence of humans that it was discovered that these birds had acquired a repertoire of words, as well as the intended whistles. The tendency for mimicked noises to be produced at first when humans are not present had been observed with a previous collection of birds, and has also been reported by Thorpe (1961, p. 119). Recording of group A was not begun early enough to discover if they behaved similarly.

Both groups produced a great amount of babbling (group B almost entirely when no humans were present) which contained the phonemes of human speech but no words. From a distance it was indistinguishable from human conversation. This behaviour is reminiscent of human pre-lingual babbling which is said to become restricted gradually to the phonemes of the adult language (Brown, 1958, pp. 199–200).

Babbling, and also native noises (shrieks and 'grackle' noises), were set going by various stimuli—human conversation, the telephone ringing, doors banging. After the end of the experiment (this applies particularly to group B), if the test whistles were played, the birds frequently joined in, on occasion joining the whistle half way through its course and finishing it. This was done exactly on pitch, as far as can be detected.

IV. DISCUSSION

The experiment showed that reinforcement did not play a part in determining which sounds the mynas would imitate. It might be argued that during the 'reinforced' sessions the birds were distracted by the sight and anticipation of food, whereas in the 'non-reinforced' sessions distractions were minimal. If this is so, it remains to be explained under what conditions a reinforcer is distracting rather than reinforcing; also it is difficult to see that any reinforcement at all could have been associated with the dusk sessions.

Since the birds in one group did their whistling (and almost all their babbling) when humans were not present, it may be that, once the response has been acquired, it has self reinforcing properties. If so, these properties were not acquired (in the case of one of the whistles) through association with primary reinforcement. The fact that the birds came to 'join in' with the test whistles suggests a kind of behavioural contagion rather than behaviour for reinforcement's sake.

The result of this experiment is disappointing. Mowrer's theory was welcome in that it showed that this particular kind of imitation learning could be explained in terms of accepted principles of learning. One is left with the unsatisfactory alternative of saying that myna birds have a tendency to mimic.

REFERENCES

BROWN, ROGER (1958). *Words and Things.* Glencoe, Ill.: Free Press.

GROSSLIGHT, J. H., HARRISON, P. C. & WEISER, C. M. (1962). Reinforcement control of vocal responses in the myna bird (*Gracula Religiosa*). *Psychol. Rec.* **12**, 193–201.

MOWRER, O. H. (1950). *Learning Theory and Personality Dynamics.* New York: Ronald Press.

MOWRER, O. H. (1960). *Learning Theory and the Symbolic Process.* New York: Wiley.

Pet Myna (1957). Fond du Lac, Wisc.: All-Pets Books Inc.

SKINNER, B. F. (1957). *Verbal Behaviour.* London: Methuen.

THORPE, W. H. (1961). *Bird-Song. Cambridge Monographs in Experimental Biology*, no. 12. Cambridge University Press.

(*Manuscript received* 24 *January* 1963)

Brit. J. Psychol. (1964), 55, 1, *pp.* 89–91
Printed in Great Britain

CONSCIOUSNESS AND BEHAVIOURISM

By J. D. KEEHN
American University of Beirut

It has been argued, especially recently by Burt, that behaviourism is untenable because the concept of consciousness cannot be denied. This argument is rejected because the functional behaviourism of B. F. Skinner not only does not reject consciousness but includes an account of how behaviour made to private, conscious, experience can be brought under public control. A simple example of how behaviour to private stimuli can be controlled through the use of collateral stimuli is described. Fuller accounts are available in the references given.

INTRODUCTION

'Behaviourism', we are told, 'is dead' (Mace, 1957). Likewise, after Burt (1962), '...behaviourism, both in its original or "naïve" form and in its later or "sophisticated" forms..., has proved untenable'. Burt goes on: '...as a basis for a general theory of human experience it (behaviourism) is hopelessly inadequate. The need to reintroduce the concept of consciousness seems inescapable.' Why *seems*? The concept of consciousness *is* inescapable, but there is no reason to deny behaviourism in all its forms because of that.

In fact there has been in existence for a number of years a variety of behaviourism that does *not* deny that men are conscious, and that actually sets out to show how functional relationships between private experiences (consciousness) and overt behaviour can develop (Skinner, 1945, 1953, 1957). Functional behaviourism, as it has been called (Keehn, 1962), is not a general theory of human experience; it is concerned with the role that experience plays in the functional control of behaviour.

FUNCTIONAL BEHAVIOURISM

The early behaviourism of Watson died (as Mace has said) in its negative aspect—the repudiation of consciousness. Psychology could not discard all references to consciousness (as Burt has said), and it was not able to dispense with introspection as one of its methods (as Woodworth has said). But Watson's early behaviourism had its positive side too—the prediction and control of behaviour; and that *the only logical and practical programme for a natural science of psychology* is behaviourism. Watson did not have the means to carry the programme out, but its feasibility has since been demonstrated time and time again.

The denial of consciousness is not so silly as it sounds. The existence of consciousness can be affirmed in the sense that we *have* experiences and that these make up the *content* of our minds. But it is not this that interests the functional behaviourist (although it is not denied that it might be of interest to somebody else—Burt, for example, and most of mankind). What matters for the behaviourist is the part played by the conscious mind in the 'causation' or control of behaviour. A functional analysis relates behaviour to the factors that control or cause it—to its antecedents—and the functional behaviourist's standpoint is that mentality is irrelevant to an analysis of this kind.

That consciousness is not the only important controlling agent of behaviour has been made most explicit by the psychoanalysts. At this point the analyst erects an elaborate mental structure between behaviour and its antecedents. To the conscious mind is added the unconscious mind and all its trappings. The functional behaviourist discards mind altogether; not, once again, as experience, but as a relevant link in a causal chain that ends in action. Putting it another way, in the classical S–O–R formulation if S, the personal history of the organism, causes O, its mental content, and there is no independent way of manipulating O, then there is no advantage in assuming the multiple linkage S–O and O–R in the explanation of behaviour. The single element S–R is enough. If the relevant functional laws are known then R can be predicted and controlled through S without knowledge of O.

But S contains more than mere reference to a contemporary here-and-now stimulus. It involves the whole history of an organism pertinent to S, R and the consequences of R to S. For example, normally when we are asked to explain a particular activity, all that is given, and expected, is a *contemporary* account: 'Why did you close the door?' 'Because it was draughty.' But the account is incomplete without the further question: 'How did you know that closing the door would stop the draught?' The answer must now be *historical*: 'Because in the past when I have closed the door the draught has stopped.' Even if there is no direct history connected with this particular door there will be *something* in the organism's history that directs its present behaviour. This is what S–R psychology means to the functional behaviourist. To say, as Burt (1962) does, that 'the subject becomes *nothing more* (my italics) than a stimulus-response machine: you put a stimulus in one of the slots, and out comes a packet of reactions' is to argue by emotional meanings (cf. Thouless, 1953). What is being said is that behaviour is lawful (and if this is not true then we are all wasting our time), and that the laws can be written in the form $R = f(S)$. The expansion $R = f(O)$ and $O = f(S)$ is rejected by the functional behaviourist on the grounds that O has no independent existence: there is no way of manipulating O except through S. If such a way ever is found *then* the behaviourist's case is lost.

PUBLIC AND PRIVATE STIMULUS CONTROL OF BEHAVIOUR

Behaviour is brought under the control of a stimulus by means of differential reinforcement (Skinner, 1938). When a response is reinforced it increases in strength, and when reinforcement is withheld the response weakens. There is a high probability that the response will be made in the presence of the stimulus and a low one that it will occur when the stimulus is absent. These facts and the technical terminology applied to them are probably too well known to need restating here, and an elementary account has been given elsewhere (Keehn, 1962).

A psychology that deals with observable responses to public stimuli faces no particular problems as a natural science, for an experimenter can see when the stimulus is present, when the behaviour occurs, and reinforce accordingly. But a conscious organism can respond to a set of private stimuli, mental processes if you like, that are not available to anyone else, and if the behaviourist is not to deny these private stimuli he must show how responses made to them can be brought under public control. Skinner claims to have given a complete account of how this is done (Skinner, 1945, 1957). The simplest case will serve as the basis of discussion.

What is necessary is to show how an organism can differentially reinforce another organism only in the presence of stimulus events not themselves directly available to the reinforcing organism. The illustration centres on verbal behaviour because this is the typical response class made to private stimuli, although other examples would do just as well. When a blind man learns the names of objects the stimulus events controlling his behaviour are tactual. They are private to the blind man and are not available to his teacher who makes *his* responses to visual cues. The blind man feels the object and names it, and the teacher reinforces the correct response by agreement. The public verbal response is made to the private 'feel', but differential reinforcement is possible because even though the blind man's experiences are private there are collateral stimuli (the sight of the object) available to the teacher. Skinner's example is dramatic because the experiences of teacher and pupil are in different sense modalities. But they need not be, for when a child learns to call a piece of furniture a table his view of the table is private and quite different from that of the adult who supplies the reinforcement. An objective *S* may be present but the experiences of the two observers are private and different.

The functional behaviourist does not need 'modern physicists like Bohr and Heisenberg' to remind him that 'all the raw material of our knowledge consists of conscious events in the lives of separate observers' for he knows that as well as any one, and finds that he can get along very nicely just the same.

REFERENCES

BURT, C. (1962). The concept of consciousness. *Brit. J. Psychol.* **53**, 329–42.

KEEHN, J. D. (1962). *The Prediction and Control of Behavior*. Beirut: Khayat's.

MACE, C. A. (1957). Behaviourism. *In* A. V. Judges (ed.), *Education and the Philosophic Mind*. London: Harrap.

SKINNER, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

SKINNER, B. F. (1945). The operational analysis of psychological terms. *Psychol. Rev.* **52**, 270–8.

SKINNER, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

SKINNER, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.

THOULESS, R. H. (1953). *Straight and Crooked Thinking*. (Rev. edn.) London: Pan Books.

(*Manuscript received* 5 *December* 1962)

Brit. J. Psychol. (1964), 55, 1, *pp.* 93–96
Printed in Great Britain

CONSCIOUSNESS AND BEHAVIOURISM

A REPLY

By CYRIL BURT
University College London

Prof. Keehn's arguments, interesting as they are, seem somewhat inconsistent with each other, and therefore by no means easy to follow. If I understand him rightly, his general position is briefly this. 'The denial of consciousness is not so silly as it sounds': admittedly we have conscious experiences, and these 'make up the contents of our minds; *but it is not this that interests the functional behaviourist*, although...it might be of interest to Burt and most of mankind'. With this last sentence I should agree; and it is precisely its restricted scope which in my view renders a purely behaviouristic treatment inadequate. My own contention is that, since 'conscious experiences' *are* of interest to most of mankind, they call urgently for systematic study on their own account—not as a substitute for, or an alternative to, the study of behaviour, but as an addition to it. And psychology is surely the appropriate branch of science to undertake such a study.

To support his paradoxical conclusions Dr Keehn relies on three main arguments.

(1) He begins by accepting the initial assumption of Watson (1919, pp. 1 f.) and Skinner (1953, pp. 10, 14 f.)—that the proper programme for psychology as a natural science is 'the prediction and the control of behaviour'. I have argued that its aim should go much further. Whether I am examining a problem-child referred to me by a teacher or magistrate, or interviewing an applicant for a post in some firm, or entertaining a guest whom I have never met before, my aim as a psychologist is not merely to *predict* the person's behaviour or to assist in its *control*, but to *understand* him. And, in the same way, I want to understand myself. Now it seems impossible to achieve any reasonable understanding either of ourselves or of other persons unless we set ourselves to study the processes and contents of consciousness by first-hand observation and analysis. The reasons for this contention I have discussed elsewhere (Burt, 1958 and refs.), and the broader view of natural science which it implies has since been independently defended by Prof. Hempel. 'Scientific inquiry', he writes, 'entails not only dependable methods of predicting and controlling events, but also a concern to *understand* the phenomena presented' (Colodny, 1962, pp. 9 f.).

In ordinary life we unhesitatingly assume, not merely that other persons are conscious, but that their consciousness is the most important thing about them. Our natural attitude (to borrow a convenient phrase from Martin Buber) is that of 'I-and-You', whereas the traditional attitude of science towards its various subjects is that of 'I-and-It'. And it is the object of behaviourism to substitute the latter for the former, even when we are dealing with human beings like ourselves. The grounds for this narrower outlook are plainly stated by Skinner (1953, p. 6). 'We cannot', he maintains, 'apply the methods of science to a subject matter which is assumed to move capriciously.' But the recent developments of quantum physics show, not only that we can do so, but that we must. Hence the argument for a deterministic treatment of human behaviour falls to the ground.

(2) I fully agree that, if overt behaviour is to be the subject of scientific research, we must assume that it is in some sense 'lawful'; and I should add that, if we are to make a similar study of conscious processes and conscious contents, we must assume that they too obey certain generalizations or 'laws'. But in my view the 'laws' must be statistical laws (i.e. probabilistic), not 'functional laws'. Dr Keehn, however, true to the postulates of a 'functional behaviourism', insists that the laws must be expressed in the simple deterministic form, $R = f(S)$, where R denotes behavioural responses, S includes not only the 'here-and-now stimulus', but also 'the organism's past history', and the symbol f, as usual, implies that for any given set of values for S a uniquely determined value can be assigned for R (if in practice we have sometimes to admit that our predicted value of R is at best only a probable value, that is due merely to the complexity of the situation and the imperfections of our knowledge, not to any essential characteristic of the so-called 'laws').

'To an analysis of this kind', as Dr Keehn rightly observes, 'mentality is irrelevant.' Accordingly, the 'expansion' of the formula adopted by the old-fashioned introspectionist, namely, $R = f(O)$ and $O = f(S)$ must be rejected, since O (the 'mental content') 'has no independent existence'. Indeed, 'R', he declares, 'can be predicted and controlled without the knowledge of O': so that, even if we were to assume that mental contents exist and can be systematically studied, such an assumption would be of no interest or help to the scientific psychologist.

Now these arbitrary simplifications seem manifestly untrue. If, for example, I know that Tommy's 'mental contents' include little or no visual imagery but do include vivid images of an auditory and kinaesthetic type, then I can predict with much greater accuracy the kind of mistakes he will probably make in his reading and his spelling, and, by adapting the methods used to teach him, I can improve both his reading and his spelling, and so control his scholastic behaviour. The same holds good in dealing with the motives, conscious or unconscious, of the young delinquent. Without a sympathetic understanding, as a glance at numerous case-histories will quickly show, the attempts of the teacher or the criminal psychologist at prediction and control must frequently be in vain.

(3) Dr Keehn's third line of argument is a modified restatement of the familiar postulate of scientific monism. 'A psychology that deals with observable responses to public stimuli', he writes, 'faces no particular problems as a natural science, for an experimenter can see when the stimulus is present.' Here the terms 'observable' or 'public' mean *open to the independent observation of the experimenter*, not just *observable by the subject himself*, even though he has been efficiently trained (this latter type of observation is, for some reason, deemed unscientific). Accordingly, if a behaviourist is to take account of conscious processes or conscious contents, which Dr Keehn calls 'private stimuli', he can only do so by dealing with them indirectly, i.e. by first reducing them to terms of data or contents that are 'public'—that is, 'observable' in his sense. To discover how precisely that is to be accomplished, we are referred to the writings of Skinner and Dr Keehn himself. From these we learn that the sole concepts that are admissible in 'natural science' are physical concepts. Thus the monism postulated turns out to be a materialistic monism of the old

Watsonian type. 'Sensations, images, and their congeries', says Skinner (1953, pp. 258, 276 f.), 'are traditionally regarded as psychic or mental events, occurring in a special world of "consciousness". The view which then emerges is almost always dualistic, and stands in the way of a unified account of nature. The contribution which a science of behaviour can make in suggesting an alternative view is one of its most important achievements....There is', he adds, 'scarcely any need to point out the disadvantages of terms which refer to supposed *non-physical* events. We cannot say why this troublesome distinction was first made. It may have been an attempt to solve certain problems.'

But surely it was an attempt, not to *solve* problems, but to *state* them. And even if the modern psychologist disclaims any interest in the 'contents of consciousness, as such, other scientists do not. Here there is room for only one commonplace example. Years ago, before suitable analgesics had been discovered, I had a couple of teeth extracted by the school dentist, and on each occasion felt a sharp sensation of pain. The pioneers of local anaesthetic would certainly not have been content to assume that all they had to consider was my *observable* behaviour, for our schoolboy code required that we should not betray our feeling by the smallest sign, and, if questioned, stoutly deny that we suffered the slightest pang. The aim of research workers like Alfred Einhorn was not to control the patient's physical behaviour, but (in his own words) 'to abolish the attendant pain'. Now the consciousness of pain cannot possibly be regarded as a 'physical event'. Hence in this case, and in many similar situations, there remains a 'dualism' which is not merely 'inescapable', but irreducible.

It is at this point that Dr Keehn's arguments reveal a certain inconsistency: for he goes on to assure us that his 'variety of behaviourism...does *not* deny that men are conscious, but actually sets out to show how functional relationships between private experiences and overt behaviour can develop'. Taken as it stands, this somewhat reluctant concession apparently implies that 'functional behaviourism' does after all study such relations as $R = f(O)$. To explain how the difficulty can be overcome, he gives an 'illustration' (drawn from Skinner's discussion of 'the private made public') 'which centres on verbal behaviour'. The blind man's tactual 'feel', we are told, is admittedly 'private'; but the experimenter can deal with it, because the blind man translates his feelings and his inferences into a 'verbal report' which is 'public'. No doubt in certain selected cases of this sort the behaviourist's account may be accepted. But in others, e.g. those mentioned in the paper criticized (Burt, 1962), such a procedure would completely miss the problem at issue. In studying differences in imagery or in investigating the 'laws' of colour-mixture the psychologist is concerned, not with the subject's verbal responses as such, but with *what* the subject reports: it is not the man's outward and visible behaviour that interests him, but the nature of his inner experience.

Finally, Dr Keehn, even when he seems prepared to admit the existence of 'mental contents', still maintains that there is no way of 'independently manipulating' them. In other words, they are at most mere epiphenomena. Such a doctrine makes nonsense of the work of the educator, the psychotherapist, the moral reformer, the legislator and the historian; for it undermines the whole notion of personal responsibility, which no one, except a psychologist or metaphysician with a preconceived theory to defend, would dream of repudiating.

Briefly then the essential difference between us appears to be this. We are both prepared to accept consciousness as an observable fact. Dr Keehn, however, regards it as of no scientific interest or importance. I on the contrary hold that, of all the empirical attributes of things that present themselves for observation and research—their position, size, and weight, their energy in its various forms (kinetic, thermal, chemical, electrical, nuclear, and so forth)—the attribute of consciousness is at once the most distinctive and (in the existing state of knowledge) the most baffling. For that reason, as it seems to me, present-day psychology, instead of evading the problem either by denying its existence or belittling its importance, should once again include within its scope the systematic study of conscious experience.

REFERENCES

BURT, C. (1958). Definition and scientific method in psychology. *Brit. J. Statist. Psychol.* **2**, 31–69.

BURT, C. (1962). The concept of consciousness. *Brit. J. Psychol.* **53**, 229–42.

COLODNY, R. G. (ed.) (1962). *Frontiers of Science and Philosophy*. Pittsburgh University Press.

SKINNER, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

WATSON, J. B. (1919). *Psychology from the Standpoint of a Behaviorist*. Philadelphia: Lippincott.

(*Manuscript received* 15 *July* 1963)

Brit. J. Psychol. (1964), 55, 1, *pp.* 97–102
Printed in Great Britain

AN ELECTRONIC STIMULUS TIMER

By R. S. RODGER*
Psychology Department
AND R. K. QUIGG
Chemistry Department
Queen's University of Belfast

The apparatus is designed to present patterns of electrically controlled stimuli such as occur in tachistoscopic work, apparent movement, conditioning schedules, etc. Its rate of operation (one cycle every 10 msec. to one cycle every 10 sec. approx.) and the duration of stimuli controlled (2 msec. to 20 sec. approx.) make it particularly suitable for perceptual work. Typically the apparatus lights neon bulbs, but a range of other stimuli are possible, particularly by using the apparatus to actuate electronic relays (thyratrons) and thus to control a much larger power source.

The design of the apparatus involves three types of cabinet. (1) A stabilized power supply (PS) operating from 200 to 250 V. (adjustable) 40–60 cyc./sec. (2) An initiation control (INIT) and (3) A stimulus control channel (CH). The PS used here can supply the operating voltages for from 1 to 4 CH as well as INIT, but a modified PS could be constructed to supply more CH.

The PS can either be purchased commercially or constructed (at rather less cost) from the circuit diagram obtainable from the first-named author. In the PS the 380 V. d.c. stabilizer delivers up to 250 mA d.c. with a hum and ripple content of less than 5 mV. and a voltage change of less than 10 mV. for a 10% change in a.c. mains input. The source impedance is less than 0·1 Ω. This stability of voltage and low source impedance is most important, since otherwise CH interaction or spurious pulses would result.

The CH each contain four logically distinct units (though not all electronically distinct). These are a pulse repetition frequency unit (prf), a time delay (del.), a stimulus duration (dur) and a constant current supply unit (cc). These units are each variable under the experimenter's control.

Each CH draws its power from PS and its prf unit generates triggering pulses at the rate of 155–0·037 per sec. (time between pulses 6 msec.–27 sec). These triggering pulses are fed via a rotary switch, with four positions (*a*, *b*, *c*, *d*), and del to dur. Del delays the passage of each triggering pulse for periods of 10 msec.–10 sec., depending on the knob setting selected by the experimenter. After leaving del, the delayed triggering pulse is fed to dur, which is a d.c. coupled flip-flop generating a voltage pulse of duration 2 msec.–20 sec. This voltage pulse switches on the cc. When cc is off, the current output is zero; when on, the current is adjustable over the range 8–25 mA d.c. and will feed a load of resistance 0–12 KΩ at 25 mA or 0–38 KΩ at 8 mA.

As well as the sockets for input from PS and output to the stimulus, there are four further sockets. two toggle switches and the rotary switch mentioned above. The four positions of the rotary switch are (*a*) off, (*b*) continuous running, (*c*) single pulses, and (*d*) external triggering. The toggle switches are for (*p*) initiating a single pulse when the rotary switch is in position (*c*) and for (*g*) switching on prf at any chosen moment of time when the rotary switch is in position (*b*). This switching on function may also be performed remotely by connecting a long cable to the co-axial socket (*s*) which is one of the four sockets on the front panel of CH. The other three co-axial sockets are (*e*) one which will receive triggering pulses from an external source (usually another CH), when the rotary switch is at (*d*), and feed them to del; (*i*) one connected directly to the input of del and (*o*) one connected directly to the output of del. From these last two sockets pulses may be taken, before or after being delayed, to trigger another CH. Taking pulses from the last-mentioned co-axial sockets does not mean that the CH providing them is in any way interrupted or interfered with; the voltage available at each of the sockets (about 20 V.) is adequate to drive up to three external CH, and since the circuits are essentially electronic triggers, small variations in voltage do not alter the timing.

---

* This author is now at the University of Sydney.

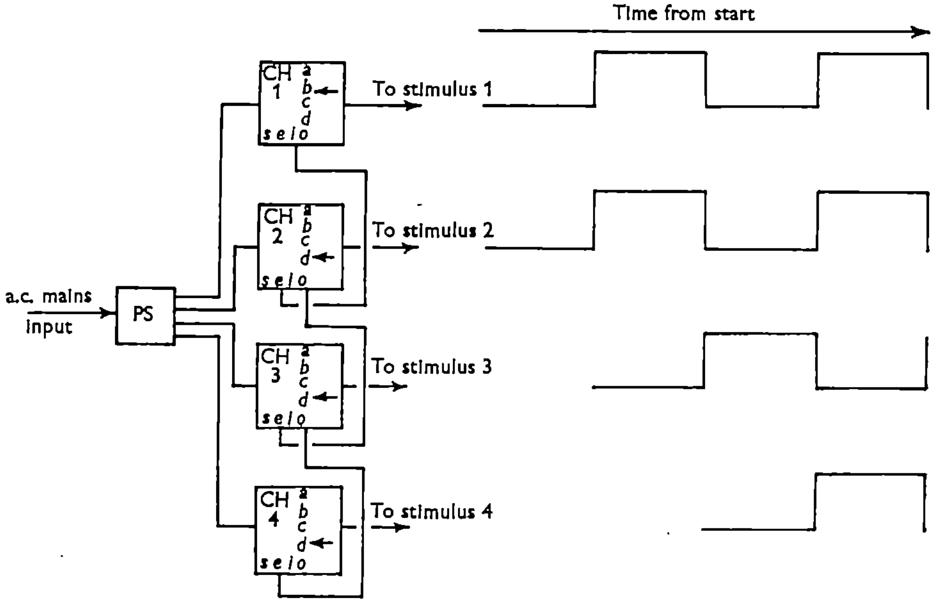
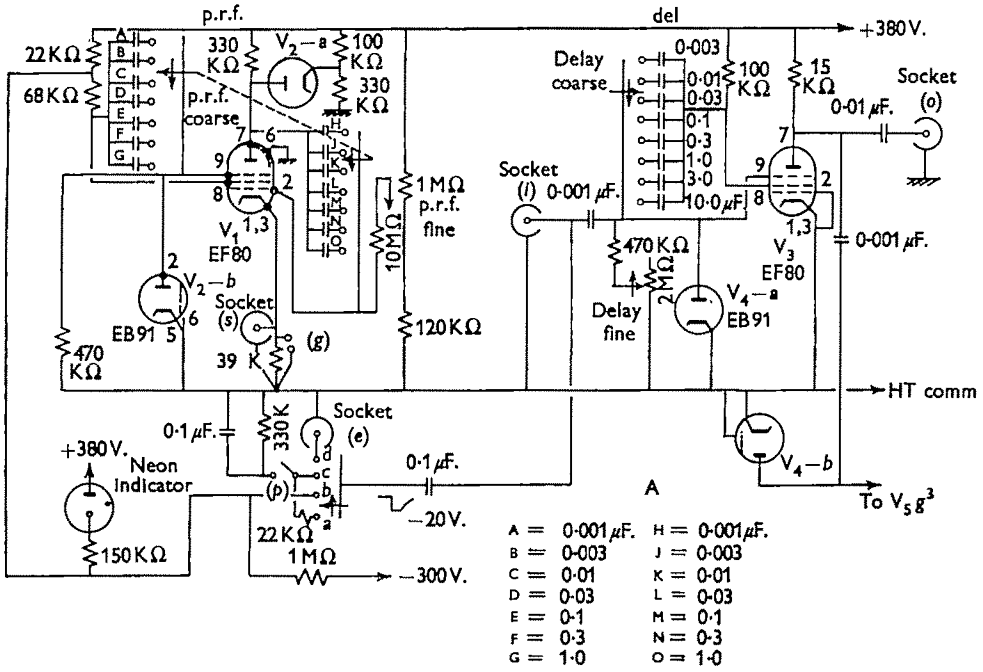


Fig. 1. CH in series.

The existence of the sockets and switches described above, and their associated electronics, means that the four CH can be operated not only as four parallel stimulus controllers but also as CH in series with one another; either two, three, four or two pairs of CH constituting a series. The block diagram in Fig. 1 illustrates a series of four CH, the prf of CH 1 acting as master with all del and dur acting as slaves. We have arranged all del, dur and cc to be equal in this example, to simplify the exposition.

Here CH 1 and 2 are firing their stimuli in synchrony while CH 3 and 4 are cascaded.

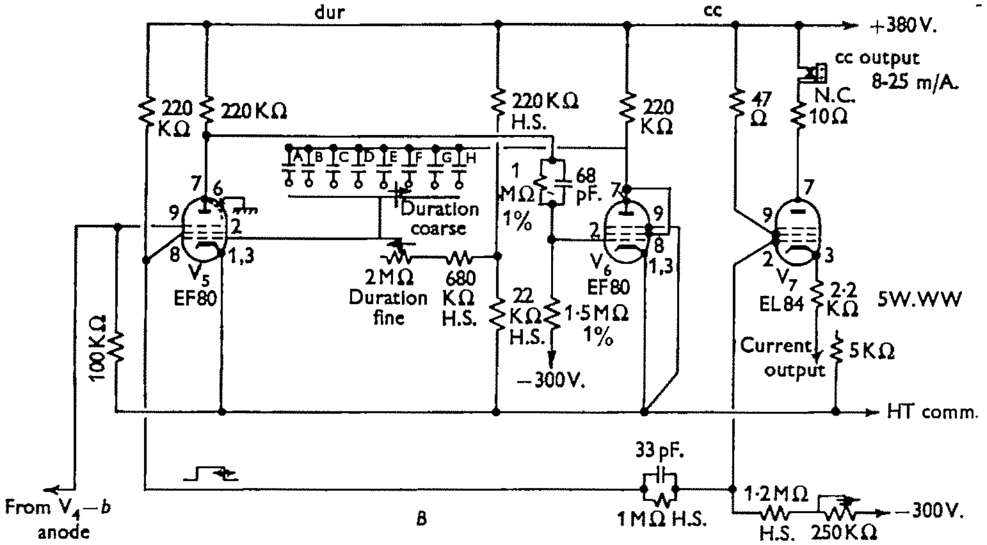
The circuit diagram of CH is shown in two parts. The prf and del units are given in Fig. 2 A while the dur and cc units are shown in Fig. 2 B.



A = 0·001µF.
B = 0·003
C = 0·01
D = 0·03
E = 0·1
F = 0·3
G = 1·0

H = 0·001µF.
J = 0·003
K = 0·01
L = 0·03
M = 0·1
N = 0·3
O = 1·0

A

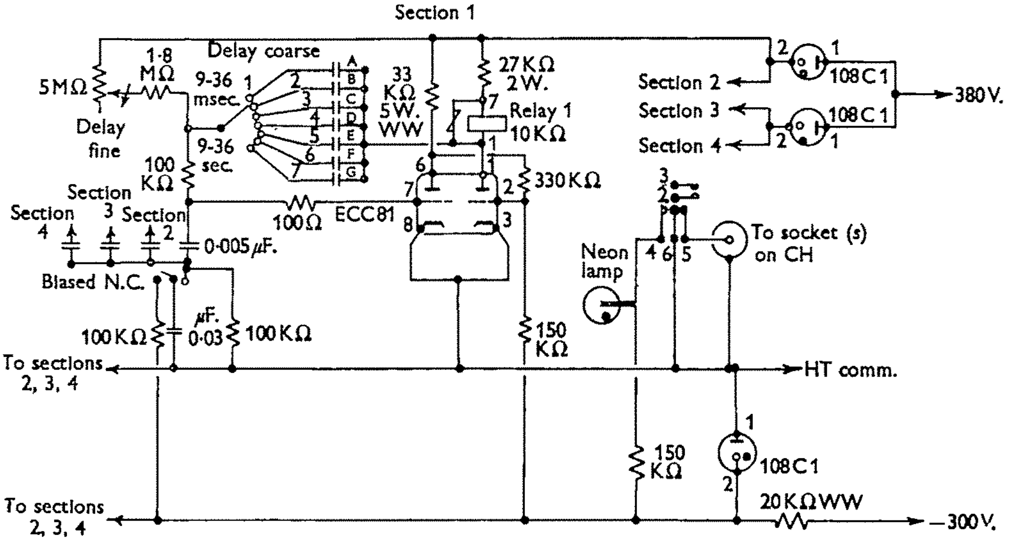


A = 0·001µF.
B = 0·003
C = 0·01
D = 0·03
E = 0·1
F = 0·3
G = 1·0
H = 3·0

B

Fig. 2. Stimulus control channel (CH).

When the CH are arranged in parallel, there is the problem of initiating all four CH so that the sequences of del/dur outputs are appropriately inter-related. One solution to this problem is to connect the (*s*) socket of each CH to a common switch which, when thrown, will initiate all CH simultaneously. There are patterns of output for which neither this simple solution nor a series arrangement is adequate. These are patterns in which the start of some CH must be delayed for a period different from the delay set thereon and also different from any combination of the delays of other CH. To solve this problem INIT was designed. This cabinet is powered from PS and has four identical sections, one for each CH. These sections are each connected to the (*s*) socket of the CH they are to control. Each section consists of a del (adjustable from 30 msec. to 30 sec.) and a relay. When INIT is switched on, the prf of all CH connected thereto ceases. No further triggering pulses are generated but any voltage pulse being generated by dur is completed. Each CH now recommences the generation of triggering pulses when switched on by its controlling section of INIT, the moment of prf switching on being determined by the initial time delay set on INIT for this section. Each CH, of course, starts with a delay set on its own cabinet. Having been thus initiated, all four CH now recycle in accordance with the knob settings on their own cabinets. The circuit diagram for one section of INIT is shown in Fig. 3.



A = 0·01 µF.
B = 0·03
C = 0·1
D = 0·3
E = 1
F = 3
G = 10

Fig. 3. Initiation control (INIT).

With a parallel arrangement there is always the possibility that the sequence of pulses on different CH will drift away from one another, because of the difficulty in setting the prf potentiometers with sufficient accuracy. One way round this difficulty is to use the finely adjustable HELI-POT potentiometers, another is to dispense with the prf units, as they are given in this paper, and instal DEKATRON counters in their place. These counters could now be fed by a single master timer (generating, say, 1000 pulses/sec.); they would be set to count off 1–1000 pulses (or more if necessary) and trigger their del circuit when their integrally accurate setting had been reached. By arranging the DEKATRONS to return automatically to zero after firing, the whole process would continually recycle. If necessary INIT could be interposed between the master timer and the DEKATRONS of any or all of the four CH, with only minor changes in the INIT design, the spare contacts already provided on the electro-mechanical high-speed relays in INIT being used for this purpose. The use of INIT is illustrated in the block diagram of Fig. 4. Again all del, dur and cc are set equal to simplify the exposition.

With fast prf or long periods of continuous firing of the apparatus, drift due to inaccurate potentiometer setting is likely to be more serious than either slow prf or short periods of firing.

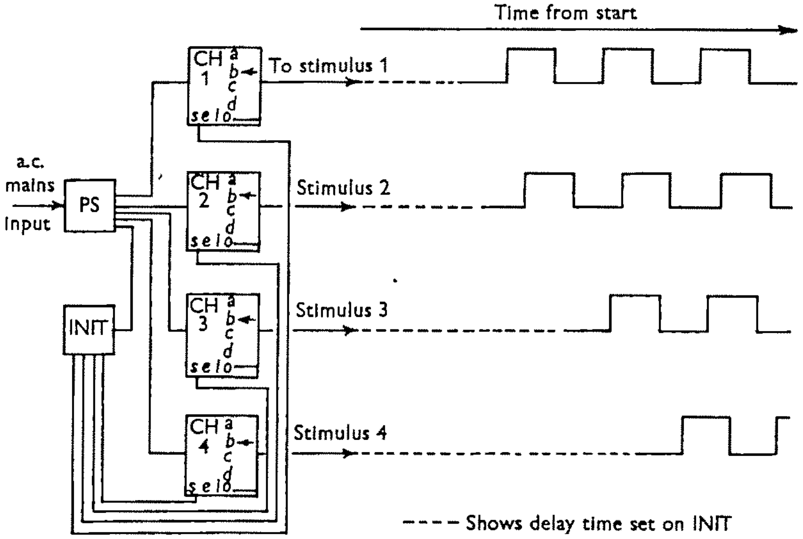


Fig. 4. CH in parallel using INIT.

The calibration of the prf units is best carried out by starting with the fastest prf's and comparing them with an accurately calibrated oscillator using a simple cathode ray oscilloscope to display a Lissajous figure. This is effective down to about 10 counts/sec. when an electronic counter and accurate stop-watch become necessary. Slow counts, from about 3 counts/sec. downwards may be carried out using only a stop-watch and counting the flashes of the neon indicator lamp built into the prf circuit. The durations of del are easily checked on an oscilloscope with a calibrated time base, for short delay times, while for longer delay times a stop-start electronic counter or a strip-chart recorder with superimposed timing marks is necessary. The dur times can be calibrated by connecting the constant current output of the CH to an electronic chronotron. The amplitude of the output can be easily measured by plugging the leads of an ordinary d.c. milliammeter into the output socket and setting dur to generate an output of about 3 sec. duration.

A full-dress calibration involves several sets of 1000 readings and occupies a number of days. Subsequent checks on the calibration are easily performed by connecting the 4 CH to a 4-channel pen-recorder, such as an e.e.g. recorder, and checking simultaneously amplitude, output duration, delay and pulse repetition frequency. The design of the apparatus is such that results depend mainly on highly stable resistors and capacitors rather than valves. Apart from catastrophic failures, the calibration, if carefully carried out, should hold indefinitely.

An initial running-in period of about 240 hr. should be given to eliminate weak or intermittent valves or other components, and to allow initial valve drift to settle. Replacement of valves in any position (except the ECC 81 in INIT and the EF 80 pentode in del) should have no effect on calibration. In the case of the ECC 81 and EF 80 mentioned, a preliminary running-in period of 10 hr. should be allowed before it is decided that a re-calibration is necessary.

All resistors used in the stimulus timer and power supply are tin-oxide 1 W. rating with a 10% tolerance unless otherwise stated on the circuit diagram. The capacitors are silvered-mica, polyester film, or paper according to size.

(*Manuscript received* 22 *October* 1962)

Brit. J. Psychol. (1964), 55, 1, *pp.* 103–120
Printed in Great Britain

PUBLICATIONS RECENTLY RECEIVED

Thought and Language. By Lev Semenovich Vygotsky. Translated by Eugenia Hanfmann and Gertrude Vakar. London and New York: J. Wiley. 1962. Pp. xxii + 168. 38*s*.

Educational Psychology in the U.S.S.R. Edited with an Introduction by Brian and Joan Simon. London: Routledge and Kegan Paul. 1963. Pp. xi + 283. 40*s*.

The Russian Tradition in Education. By Nicholas Hans. London: Routledge and Kegan Paul. 1963. Pp. vii + 196. 30*s*.

It is in some respects surprising that a scientific book published as long ago as 1934 should still have in 1963 more than historical interest. Perhaps the growing interest in thought and speech and their interrelations has made western psychologists increasingly dissatisfied with the explanatory value of speculative chains of fictitious *S–R* units, involving neither receptors nor effectors—a paradoxical invention of the *S–R* school and an implicit admission of defeat—and a book such as Vygotsky's *Thought and Language* which regards thought and speech as distinct but interrelated and therefore recognizes and discusses the problems of their interrelationship, which refuses to deal with either in terms of reflexology, keeps the discussion on a factual and often an experimental level, presents a relatively novel point of view. Of course many in the West have never been reflexologists—Western psychology is too diversified for anything like uniformity—but it would not be untrue to say that the majority, especially in America, have been and perhaps still are reflexologists.

Reflexology was and still is psychological orthodoxy in Russia, but Vygotsky's book was published before the penalties for unorthodoxy became prohibitive. It was banned in 1936 and not republished till 1956 when the situation had eased considerably, so that psychologists could work safely in such fields as those of thought and speech, provided that orthodoxy was not openly flouted. They have been greatly aided in this by Pavlov's great authority even after his death; for he had insisted that the problems of thought and speech could not be solved in terms of his great creation, the first signal system—a second signal system was necessary. But as the theory of the second signal system does not exist in the sense that that of the first signal system does, being little more than an aspiration, 'music of the future', the effect has been to give psychologists greater freedom and security.

Vygotsky's book must be judged primarily as a theoretical work; that is, rather for the ideas and points of view it puts forward than for the experimental findings it describes. (The latter are in fact described very succinctly, perhaps too much so for appraisal; the only experiment described sufficiently fully, that of Zhif (pp. 105ff.), is defective in more than one respect.)

Vygotsky's views are partly a critique, partly a development of those of Ach and of the early Piaget, especially the latter. The book is enriched by a short essay by Piaget relating Vygotsky's theories and criticisms to his own earlier and later theories. (Unfortunately the pamphlet containing this essay is not bound up with Vygotsky's book and is liable to be lost.)

Vygotsky died in 1934. Two years later the great frost began with the prohibition of psychometry. Here too his views were at variance with the official position. True he criticizes current intelligence tests, but with a view to improving, not abolishing them. His early death—he was only 38—was a great loss to psychology, but in view of his unorthodox and original mind it is doubtful whether, even if he had survived, he could have made the great contribution to psychology of which he was capable. This book (which incidentally is presented in a considerably abridged form) makes one realize how great that could have been.

Educational Psychology in the U.S.S.R. consists of a collection of articles, often abridged, on various topics of interest and importance to educational psychologists. For the Russians it is clear that schoolchildren's learning is not only the chief but almost the only topic of educational psychology. Compared with the range of topics discussed in almost any English or American text the omissions are striking. Striking too is the uniformity of the theoretical standpoints of the contributors.

How children learn to read and to solve arithmetical problems, what sort of ideas they have in physics, such are examples of the very down-to-earth problems, tackled experimentally in schoolroom settings, that are described in Part III of this book. Part II consists of an abridged translation of an article by Bogoiavlenski and Menchinskaia on learning, published in 1960 in Volume II of *Psychological Science in the U.S.S.R.* This gives the best account available in English of Russian work in this field. Part I is more theoretical and less satisfying, partly because of too great compression and repetition, on the one hand, and partly because of being too polemic and apologetic, on the other; it could very well have been given less space—it is a quarter of the book—in favour of less abridgement elsewhere. The editors' introduction describes the organization of psychology, psychologists, and psychological departments in the U.S.S.R.; the degree of centralization and state control is impressive and probably accounts for the far greater uniformity of Russian psychology as compared with that of the West. There is no doubt that most readers of this book will gain a much clearer picture of Russian educational psychology than they had before, even if scientifically they conclude that they have gained less than they had hoped.

The Russian Tradition in Education is more of interest to the educational historian than to the psychologist. Its chief aim is to show that a Russian tradition in education had developed by 1917 and has persisted since. 'Throughout the history of Russian education its national character was quite evident...Russian it was and Russian it remains.' The greater part of the book—150 pages—is concerned with establishing what the tradition was before 1917, the remaining 40 pages with showing its persistence since. The tradition in question is rather one of doctrine than of practice and is defined, so far as definition is possible, by considering the views of outstanding educationists.

H. E. O. JAMES

The Psychology of Time. By PAUL FRAISSE. Translated by JENNIFER LEITH. New York: Evanston; London: Harper and Row, 1963. Pp. vi+343. 41*s.* 6*d.*

British psychologists will wholeheartedly welcome this admirable monograph on Time, the first that has appeared in English since Mary Sturt's book was published a generation ago. Professor Fraisse's volume is, however, far superior to its lonely predecessor, for it displays a full measure of philosophic breadth, extensive reading and critical acumen. His book is clearly not one that could be written over a week-end. It is the fruit of years of devoted and dedicated study during which he not only assimilated a vast body of knowledge but also contributed more richly and diversely to his subject than anyone else.

Although he clearly owes much to the work of Henri Piéron, whom he succeeded at the Sorbonne, and whose pioneer experiments on time he has greatly developed and elaborated, his own work shows more flexibility than Piéron's and it is happily free from anti-Bergsonian or, indeed, any other bias. It is significant that Fraisse, above all an experimentalist, is as much at home in the world of Janet as in that of Piéron, and one of the many virtues of the present work is that it brings home to the English reader the force and originality of the ideas of Janet, who has received scant attention from British psychologists. There is also an exciting analysis of the Piaget–Fraisse controversy over the psychogenesis of time.

Fraisse's principal aim is to investigate the manifold ways in which man 'adapts to the temporal conditions of his existence'. With this goal in mind he organizes his material in terms of three levels of adaptation, viz. (i) conditioning to change; (ii) perception of change; and (iii) control over change. One advantage of this approach is that it carries the discussion in language intelligible to biologists and physiologists as well as to philosophers of all schools. There is a full bibliography of some 600 items which the student of time, from whatever point of view, will find of great value.

In this all-too-brief review, I can only emphazise that here is a book from which all psychologists can profit, for it is the most thorough-going synthesis yet attempted of our knowledge of a theme of fundamental importance: time in human life.

JOHN COHEN

The Philosophy of Mind. Edited by V. C. Chappell. New York: Prentice Hall (Spectrum Paperbacks). 1962. Pp. ix + 178. $2.25.

For most psychologists, I suppose, 'behaviourism' still suggests J. B. Watson and all that. Yet today there are two quite distinct, and even antagonistic, versions of behaviourism. Let us call them the left wing and the right wing. Left wing behaviourism is reductive and mechanistic and bases its claims on science; right wing behaviourism is antireductive and anti-deterministic and bases its claims on semantics. The present volume is a compilation of published papers all of which, with two striking exceptions, belong more or less on the right wing.

The reader would do well to start with D. W. Hamlyn's article, which discusses behaviour in the light of Aristotle's distinction between activity and movement. In psychology we are concerned with activities rather than movements and since only movements can properly be said to have causes Hamlyn concludes that a strictly scientific psychology is an impossibility. But the true father of right wing behaviourism is Wittgenstein rather than Aristotle, so the reader should turn next to Norman Malcolm's review of the *Philosophical Investigations*. Malcolm, it is true, denies that Wittgenstein was a behaviourist and attacks Strawson for saying so and, since Wittgenstein's mode of philosophizing had its built-in-defences against being pinned with any label, Malcolm can of course get away with it; but can we really doubt that, for example, Ryle's *Concept of Mind*, which is here discussed by John Wisdom, is a legitimate child of the *Philosophical Investigations*?

Other contributions which should interest psychologists are: P. F. Strawson on 'Persons', in which he attempts a monistic, as opposed to the Cartesian-dualistic, interpretation of the concept of a person; Errol Bedford on 'Emotions' in which he proposes an analysis in terms of the social aspects of emotional behaviour in place of one in terms of inner feelings; and B. A. Farrell on 'Experience' (from his 1950 article in *Mind*) in which he argues that the way to get rid of the notion of experience and mental events is 'by realizing it is just foolish to suppose that there are, or are not, any such things'.

The two striking exceptions I mentioned are: U. T. Place on 'Is Consciousness a Brain Process? (which first appeared in this *Journal* in 1956) and J. J. C. Smart on 'Sensations and Brain Processes'. These authors use the apparatus of linguistic philosophy to establish a materialist thesis, namely that it *is* possible to identify mental events with brain processes. Their reasoning is highly ingenious and should not be missed, but its validity depends in the last resort on one's agreeing that what Place calls 'the phenomenological fallacy' is indeed a fallacy. J. BELOFF

Psychology through Experiment. Edited by George Humphrey. London: Methuen. 1963. Pp. vii + 307. 30*s*.

The purpose of this volume is partly to act as a companion to work in university practical courses, but the book is also intended to appeal to non-specialists as an introduction to the methods of research employed by psychologists. Correspondingly, there are two introductions by the Editor, one being for humanists and one for professional psychologists. The book itself consists of seven chapters each by a different person and each discussing a different field of psychology. The areas covered are sensory processes, perception, remembering, learning, motor performance, thinking, and animal experimentation.

There is a difficulty for any editor of such a book in that he must steer between the Scylla of imposing a dull uniformity on his contributors, and the Charybdis of letting them become so individual as to leave no common thread to the book. In this case, 'the policy has been to interfere as little as possible with what they have written'. The effect of this has been to produce a book which is lively and stimulating, but in which the intention of the different chapters appears to be somewhat different. Thus R. L. Gregory in discussing sensory processes gives primarily a chapter containing information about the eye and ear, while mentioning from time to time cheap and ingenious demonstrations which could be used by a lecturer to add variety to his teaching. B. M. Foss, on the other hand, devotes his chapter on thinking to three experiments which could be conducted by the reader himself and in which details such as the experimental design, the exact instructions to be given to the subjects, and so on are specified, while the theoretical content of the chapter is kept to a bare minimum. A third method of approach is that of

J. A. Deutsch, whose chapter on animal research contains some general and very practical hints on the handling of rats, and descriptions of three experiments which he has himself published, together with the theoretical arguments to which they are relevant. Each of the other authors similarly has his own preferred pattern of presentation.

Because of this variety, the book does not provide a ready-made practical course, nor a set of demonstrations, nor an assemblage of experimental facts about psychology. Parts of it, however, are of value to anyone who is interested in each of these purposes. It is rather a quarry from which the university teacher may extract ideas, than a self-service shop where he can be provided with a ready-made practical course. There is of course an abiding stream in British academic life which holds that the object of university education is to teach people to think for themselves, and accordingly that one should not present tightly organized courses on the American model. This book should appeal to those who support this point of view.

DONALD E. BROADBENT

The Prediction and Control of Behavior. By J. D. KEEHN. Beirut: Khayat's. London: Constable. 1963. Pp. 268. 45*s*.

Introduction to Method in Psychology. By W. M. O'NEIL. Second edition. Melbourne: University Press; London: Cambridge University Press. 1963. Pp. xii + 178. 30*s*.

Dr Keehn subtitles this book 'A shorter introduction to psychology'. This book appears too complex, however, to prove worth while for the undergraduate without help from a number of other texts. The articles and books of B. F. Skinner should prove useful here.

Behaviourism, a *functional* behaviourism, is proposed as the system most encouraging to research in psychology, and to explanation of the psychological facts present today. Functional behaviourism does not exclude consciousness from psychology and is thus broader than the earlier, *analytic* behaviourism (e.g. of J. B. Watson). Most analyses of human behaviour involve correlations between uncontrolled responses, e.g. between scores on intelligence tests and academic or occupational success. Experiments with animals often deal with responses controlled by known stimuli, e.g. rats trained to press a bar in their cage on presentation of a light or tone. This training requires some kind of reinforcement; often pellets of food are presented upon the correct responses. Behaviour, as influenced by learning, personality, and motivation is well covered by this book.

The first edition of Prof. O'Neil's book (1957) has been reviewed in this *Journal* by Drever (August 1958). In the light of criticisms from this (and other) reviews the author has made several minor revisions for the 1963 edition, which appears slightly (albeit not substantially) more systematic. Graduate research students may find this book useful; it might however confuse the undergraduate.

H. WALLERSTEIN

Learning and Instinct in Animals. By W. H. THORPE. Second edition. London: Methuen. 1963. Pp. x + 558. 63*s*.

A Study in Behaviour: Principles of Ethology and Behavioural Physiology, Displayed Mainly in the Rat. By S. A. BARNETT. London: Methuen. 1963. Pp. xvi + 288. 45*s*.

This revision of Thorpe's *Learning and Instinct in Animals* is most welcome although, in fact, the first edition remains unrivalled in its field both for the brilliance of its synthesis and the quantity of comparative information to be found in it. Nevertheless, there has been an extremely rapid growth of knowledge in the fields covered by the book, which would have inevitably dated it prematurely. The revision which Thorpe now offers us is particularly valuable since it incorporates those discoveries in the fields of genetics and the physiology of learning with which students both of zoology and psychology must become acquainted. The revision has been extensive, but not so great as to alter the basic orientation of the book which gave it its unique value. Both the research worker and the student will continue to find this book indispensable for many years to come.

Barnett's study of the rat is another book that should find a warm welcome. For the psychology student its particular value lies in the well-balanced account which it gives of all aspects of that very familiar rodent's behaviour. It should serve to correct what has often been, in the past, an unfortunate one-sidedness in the experimenter's knowledge. Students will find the clear, comprehensive and well-documented account of experimental work on the rat most useful, and it is doubtful if many research workers will be able to read it without finding something new in the way of information or insight. The level of difficulty is at times a trifle uneven; e.g. 'stimulus' is carefully explained while 'phenotype' slips by without comment—but for the most part the book is commendably clear, and free of unnecessary jargon.

The ethological orientation of both authors reminds us once again how great is the debt of comparative psychology to that science; the merits or limitations of much experimental work become clear when it is set in a context which takes account not only of the whole range of the animal's behaviour, but also of that of its nearest relations. It is this overall grasp that makes Thorpe's work of particular importance; it is both genuinely comparative and at the same time free from the conceptual limitations which the rather narrower behaviourism that Barnett advocates may impose. So that, while Barnett's study will fill an empty space on the shelves of most psychologists who work with rats, Thorpe's will, I think, provide a theoretical orientation productive of research. Indeed it seems not unlikely that Dr Thorpe may be providing some of the stimulus to that advance in knowledge which has made it necessary for him to revise his own contribution so soon.

M. M. LAWLOR

Integrating Principles of Social Psychology. By JOSEPH B. COOPER and JAMES L. MCGAUGH. Cambridge, Mass.: Schenkman. 1963. Pp. xi+320. $7.50.

At a time when it is becoming increasingly difficult to find the text of many American text-books among the proliferation of tables, figures, guides, glossaries, and boxes, this European-type book would appear to be welcome. The authors, who teach at San Jose State College and Oregon University respectively, present the conventional topics such as socialization and communication but also promise infra-human behaviour analogies and a discussion of the philosophical and psychological theories of human nature. The material is examined in the light of their three integrating perspectives; viz. that the study of social behaviour must focus on the social individual; that the facts provide strong support for a cognitive-theoretical position; and that social behaviour must be viewed as a constant process of adaptation and hence that comparative and ethological information is vital.

In fact we are presented with a literary and often quaint account. The general level of discussion is elementary and the descriptions of phenomena are based heavily on those of established authorities, e.g. Allport on prejudice, Gibb on leadership, and Krech and Crutchfield on attitudes. The animal data is often anecdotal though novel. Cooper himself has studied the unusual dominance relationship among captive African lions, evinced in the multiple-chain tail chewing ceremony. The almost Victorian atmosphere is heightened by the detailed consideration of the data from feral children and the discussion of theories of the origin of language. With regard to the former, the new evidence of Ogburn and Bose casting serious doubt on the whole episode is not mentioned. In the language section the onomatopoetic theory is duly dismissed, it would hardly appear to warrant space in a chapter that has little social import and does not consider recent sociological work, such as that of Basil Bernstein, at all.

Neither as a factual text-book, nor as a book of ideas does it give a fair picture of what social psychology is about today. It is over-priced.

HALLA BELOFF

The Developmental Psychology of Jean Piaget. By JOHN H. FLAVELL. Princeton N.J. and London: Van Nostrand. 1963. Pp. xvi+472. 68*s*.

The print on the dust cover suggests that the 'book offers a detailed and integrated summary of the work of Jean Piaget'. After a short introductory chapter in which the author gives a brief historical sketch of Piaget and his work, the volume is divided into three sections. Part I deals with the theoretical hypotheses and opens with a discussion of such basic terms as *structure*, *assimilation*, *operation*, and the like. This is followed by a general outline of the characteristics of the sensori-motor period, and of the special features that emerge, such as the permanency of the object, the first notions of space, time, causality; also of imitation and play. The latter part of Chapter 4 deals with the nature of representation and the fundamental importance of symbolic functioning, while Chapter 5 is devoted to the nature of concrete operations. Formal operational thought is dealt with in Chapter 6, but curiously enough, Flavell has devoted the last dozen pages of this chapter to Piaget's work on perception. In Part II the author deals first with the experimental studies of Piaget and his co-workers, including studies from both his early and later work, and then with the more recent validation studies carried out by others. The final part contains Flavell's own evaluation of the over-all Piagetian system, together with criticisms and complaints regarding the experimental work of the Geneva school.

The author is to be congratulated on writing a book that gives so extensive a coverage of both the theory and the experimental work of Piaget. The text is clear in the main, it is well documented with earlier and up-to-date references (to the time of going to press), and it attempts a sensible evaluation of the Piagetian system. There are a few blemishes. Flavell sometimes repeats himself unnecessarily, the reports of a few validation studies are sketchy, discussions of the theoretical and experimental aspects of perception appear in odd places, and occasionally he misrepresents, as, for example, on p. 369 where he plainly implies that Dienes's work is explicitly grounded in Piaget's theory. But when Flavell's writing is not clear, he is usually reflecting the underlying vagueness in the foundations of some aspects of Piaget's system. However, taken as a whole, this will be a most useful text for more advanced students of psychology and education.

K. LOVELL

Infant Care in an Urban Community. By JOHN and ELIZABETH NEWSON. London: Allen and Unwin. 1963. Pp. 268. 42*s*.

Six Cultures: Studies of Child Rearing. Edited by BEATRICE B. WHITING. New York and London: Wiley. 1963. Pp. 1017. 95*s*.

Both of these books present material obtained in the first stages of longitudinal researches. In both undertakings the purpose was to investigate and compare the effects of different climates of social experience on children's personalities.

Drs John and Elizabeth Newson, research psychologists at Nottingham University, set aside a voluminous literature of conflicting theories and hypotheses concerning child care and collected afresh some of the data which they considered basic to their task. Their efforts have resulted in a well-planned study of 709 normal 1-year-old children in ordinary families living in a large Midland city and representing all social grades from professional to unskilled. The informants were mothers of children who were within 2 weeks of their first birthdays. The technique used, a 'guided interview' based on a well-structured questionnaire supplemented by tape recordings, appears to have been highly successful in eliciting a surprising amount of information in the course of relatively short interviews. All of the results are interesting; some are surprising, while the implications of one or two may possibly be found unpalatable in some quarters. Throughout the text, a number of tables, brief but adequate, help to summarize main findings at various stages. There are four appendices and a useful list of references to books and articles published within the past ten years. The authors are to be commended for their clear, jargon-free and eminently readable style.

The collection of six self-contained monographs, edited by Beatrice B. Whiting of Harvard University Human Development Laboratory and written by twelve contributors, constitutes the first volume of a planned series of cross-cultural studies carried out in six different parts of the world. The field work was undertaken by six teams, each of two social scientists, from the Universities of Cornell, Harvard and Yale. The presentation of material follows a similar plan in all the contributions, so facilitating comparison, and emphasis throughout is on the cultural backgrounds of children, aged between 3 and 10 years, in the chosen areas. Each monograph provides a wealth of ethnographic material with illustrative photographs. A short bibliography is appended to each, though in none of these could reference be found to even one publication having obvious connexions with child rearing. A more apt title for this particular volume might possibly have been 'Six Cultural Backgrounds to Child Rearing'.

M. I. DUNSDON

The Housing Environment and Family Life. By D. M. Wilner, R. P. Walkley, T. C. Pinkerton and M. Tayback. Baltimore: Johns Hopkins Press; London: Oxford University Press. 1962. Pp. xxii + 338. 60*s*.

The major points of interest in this book are the simplicity of the design of the research which it reports and the significance of the choice of subject-matter in relation to the cost of the research. The inquiry was concerned with the effects of improved housing on physical health and social psychological adjustment. It was conducted in Baltimore by the School of Public Health of Johns Hopkins University and dealt with two carefully matched groups of Negro residents, one of which moved into a new housing scheme in the city whilst the other stayed in the slums. They were interviewed before moving and ten times during the 3 years after by a number of unnamed interviewers using questionnaires dealing with morbidity (each time), social relations and psychological state (three times), and housing (twice). School records were also compared. The voluminous results are reported with simplicity corresponding to the controlled 'before–after' design.

This programme set out to give unequivocal answers and does so within limits. One of these limits was that the respondents were exclusively housewives, which is such a common limitation in social research that one might legitimately have hoped for something different in such a comprehensive and expensive inquiry (5 years planning, 3 years field work, $500,000). Although the researchers clearly have a first rate local knowledge, as shown in occasional remarks and in the make-up of the measuring scales, it would have helped the reader to evaluate the significance and generality of the work if the writers had given a more vivid supporting picture of the daily lives of the people concerned. Apart from these points, the whole style of the research and the report is sophisticated, sensible and thorough.

The changes demonstrated in the test group are modest but difficult to summarize—mainly some improvements in health, especially for children, some increase in neighbourliness and some improvement in felt social status and morale. The value of the work lies in the demonstration that these effects are specifically related to changes in housing. The research was clearly directed towards the consideration of some common bones of contention about public housing policies and their effects, and some of its preambles and some odd remarks seem to be carefully addressed to a local audience. The book is of considerable general importance, for it is a rare example of a very simple exercise conducted on a grand scale in a real life situation. It has produced reliable if modest results and shows how much it can cost to do just this.

G. J. HUTTON

Psychology in Relation to Medicine. By R. M. Mowbray and T. Ferguson Rodger. Edinburgh: Livingstone. 1963. Pp. vii + 399. 36*s*.

All interested in medical education agree that psychology should be taught to pre-clinical medical students and that so far such teaching has been very inadequate in most British medical schools. The majority of psychiatrists would probably agree that the teaching should consist of a mixture of experimental, dynamic, clinical, medical and social psychology which would prepare the student to deal with the doctor–patient relationship and form a basis for his future training in clinical psychiatry. However, up till now there has been no psychology text-book that caters for the medical student's needs. Many of us were hoping that this book would be suitable for this purpose, but unfortunately it is not.

The book is based on lectures given to the pre-clinical students in the Glasgow Medical School. As is well known, it is difficult to turn lecture notes into a book, because one tends to forget, as these authors have done, that what seems witty, amusing and instructive in a lecture may appear facetious and banal in cold print. The style of this book, therefore, leaves very much to be desired. There are too many topical and parochial allusions, there is an excessive use of exclamation marks and the abundant scotticisms will appear strange to the English reader.

The over-all impression is that this book has been written mainly by the clinical psychologist of the team and that the psychiatrist has added a few clinical details. Thus, where experimental work in psychiatry is well known to psychiatrists, such as Payne's work on schizophrenic thought disorder, it is mentioned, but where such work is less accessible, for example, Talland's work on the Korsakoff psychosis, there is no reference to it. Some omissions are even more striking; for example, there is no adequate discussion of the psychological, physiological and clinical aspects of disordered consciousness, which is an important problem for the practising clinician. Here and there the authors appear to have deviated from generally accepted opinions. Thus, in their discussion of hallucinations they write, 'The late-night reader who has just finished a tale by Edgar Allen Poe will have a set to misperceive a creaking floorboard or billowing curtain in a sinister fashion. This explanation accounts for hallucinations as faulty interpretations of actually present stimuli.' This is a remarkable statement, when one remembers that in 1832 Esquirol separated illusions from hallucinations and defined them as misrepresentations of actual stimuli.

Quite apart from these minor faults, the book is far too discursive and fails to deal adequately with many topics, although there is sufficient space to do so. It is unfortunate that this book fails in its purpose and cannot be recommended to the medical student.

FRANK FISH

A Study of Brief Psychotherapy. By D. H. MALAN. London: Tavistock. 1963. Pp. xiv + 312. 35*s.*

An Introduction to Psychotherapy. By SIDNEY TARACHOW. New York: International Universities Press; London: Bailey Bros. and Swinfen. 1963. Pp. viii + 376. 68*s.*

This brace of books provides an interesting contrast in current thinking on psychotherapy. Malan, under Balint's guidance, has set out to investigate a particular problem within the framework of conventional psychoanalysis and has reformulated a bit of the orthodoxy. Tarachow, on his own, offers a ponderous personal version of psychoanalytic traditions.

Malan examines whether psychoanalytically based treatment of patients with severe psychoneurotic illnesses can be successfully accomplished in between ten and forty sessions. The case material is derived from 21 adult patients at the Tavistock Clinic, treated by 7 different therapists, whose cases were discussed at weekly meetings led by Dr Balint. The conclusion is that, provided (*a*) the therapy is active, (*b*) the patient is strongly motivated and able to respond to interpretations, and (*c*) the patient and therapist are willing to become deeply involved, the brief form of treatment is as effective as the longer. This obviously begs the larger question of the efficacy of psychoanalytic treatment in general, for at the end of it all, ten of the group have a zero rating on a four-point scale of improvement. Malan's method is, however, more interesting than his conclusion. He discusses and tackles many of the awesome problems of method at the level of clinical judgement and commands respect for his candour. Psychologists who have dreamed of tackling the problem of evaluating psychotherapy other than by counting heads should read Malan, if only to learn how much sheer enthusiasm will be required of them. This enthusiasm is conveyed in the book by a persistent use of the first person and a somewhat discursive style.

Tarachow's style is first person oracular. It is monologue disguised as dialogue, for the book purports to present transcripts of supervisory sessions between the author and an unnamed hero of a resident who manages to steer the author through a long list of topics—the cardinal issues of psychotherapeutic technique (Tarachovian), active therapy in the hospital setting and how it is complicated by administration and discipline, and the place of values in treatment. The style is contrived and tedious and many of the issues are trivial. This book is not for psychologists who wish to learn about psychotherapy, but it could serve as a source-book for those who want to develop or to maintain prejudices against this field.

R. M. MOWBRAY

Schizophrenia: Conditional Reflex Studies. By CHRISTIAN ASTRUP. Springfield, Ill.: Thomas. 1962. Pp. xx + 345. $13.75.

In this monograph, one of the series *American Lectures in Objective Psychiatry*, Dr Astrup sets out to review the experimental studies of schizophrenia developed in the context of Pavlovian theory. The author has himself worked in psychiatric clinics in Eastern Europe and in Russia and his emphasis throughout the book is on studies which have not hitherto been reported in this country. In this connexion he provides a most valuable index of 1105 references, the majority of which are Russian. Apart from reviewing the work of other people engaged in what the author refers to as reflexological psychiatry, Dr Astrup also reports on his own studies of 306 chronic and 178 acute cases of schizophrenia. The book falls naturally into four main sections, the first of which contains a full description of the experimental methods used in reflex studies of the higher nervous activity in human subjects. The second section summarizes previously published reflex studies relevant to psychiatry and this is followed by a long clinical section in which the author outlines the Kleist–Leonhard system of classification of the schizophrenias, which he uses in his own studies to subdivide clinically his chronic psychotic group. In the final section, Dr Astrup reports his experimental findings and their correlation with the clinical subtype of schizophrenia.

Although one welcomes this attempt to present a rigorously experimental approach to the pathophysiology of schizophrenia, the results are often unconvincing and somewhat ambiguous. The analysis of the experimental findings provides the general conclusion that, 'the more the patients are disturbed and deteriorated, the greater is the impairment of the higher nervous activity'. As one also suspects that some of the experimental techniques used in these tests merely measure the subject's capacity to co-operate in the task, rather than direct disturbance of the nervous system, this finding is perhaps not unexpected. One's doubts about the validity of the experimental measures of higher nervous activity are also heightened by the author's finding that when the schizophrenic patients are later re-examined on the reflexological tests, 'the experimental data, contrary to expectations, has given as much improvement as unimprovement of disturbances of the higher nervous activity'. Dr Astrup explains this failure of the tests to correlate with changes in the clinical state, by suggesting that the use of ataraxic drugs may have counteracted the course of the schizophrenic deterioration. The author's main aim was to seek correlation between his experimental data and different forms of schizophrenia, and the Kleist–Leonhard system of classification is chosen as the most reliable method of clinical subdivision. After producing initial evidence to assert the comparative objectivity of the clinical classification, Dr Astrup later reaches the opinion that the differences between the clinical subgroups are not very great and that it might have been more practical to use the wider categories of slightly and severely deteriorated paraphrenia, hebephrenia and catatonia. Towards the end of the book his own doubts about the reliability of the clinical classification used become more evident, and he closes his reports with the comment, 'One entertains the idea that the clinical differences have been over-estimated compared with resemblances'. It is, however, only too easy to be critical of this attempt to introduce objective measurement into psychiatric classification and the author himself is well aware of the wide gap that exists between the aims and the actual knowledge of experimental psychiatry. It may come as a surprise to many readers in this country to find that such a solid body of work already exists in this direction and one hopes it may stimulate some readers to pursue this line of approach which must inevitably add to our very scant understanding of the basic psychophysiological disturbances associated with schizophrenia.

ANDREW MCGHIE

La Cécité Psychique. Etude Critique de la Notion D'Agnosie. By HENRY HÉCAEN and RENE ANGELERGUES. Paris: Masson. 1963. Pp. 202. 35F.

This analysis of 'psychic blindness' is based on clinical study of 415 cases with posterior cerebral hemisphere lesions, of which 102 presented higher visual disabilities of one form or another.

The principal forms of disability recognized by the authors are as follows: (1) visual object agnosia, or inability to recognize objects by sight; (2) unilateral spatial agnosia, or neglect of one half of extra-personal visual space; (3) more complex forms of spatial agnosia, including defects of spatial orientation or topographical memory; (4) defects of physiognomic recognition (prosopagnosia); (5) agnosia for colours; and (6) defects in pictorial perception and recognition (simultanagnosia).

The incidence and frequency of these various syndromes are considered. It is found that (1) is very rare, and is almost always associated with bilateral occipital lobe lesions; (2) and (3) are relatively common, and are linked predominantly with right hemisphere lesions—as is (4), though it is a good deal less common. (5) is rather rare and is found almost exclusively with left occipital lesions, while (6) is not specifically related to lesions of either hemisphere. It is pointed out that the frequency of all forms of visual agnosia is increased if the lesion is bilateral.

The authors have some interesting things to say about the relative importance of locus and extent of lesion in the causation of visual agnosia. Whereas a 'mass action' effect undoubtedly plays some part in its origin, especially in the case of right hemisphere lesions, localization is also an important consideration. For example, colour agnosia is narrowly related to lesions of the left occipital lobe. In general, the authors consider that defects arising from left hemisphere lesions appear to have a more restricted anatomical focus than is the case with defects arising from right hemisphere lesions. This is in agreement with the view expressed by Teuber and his co-workers in their recent study of cortical sensory loss.

Although this study is entitled 'critical', it cannot be said that the authors adopt a particularly critical attitude towards theoretical issues. Nor is there any real attempt to analyse further the very interesting psychological defects revealed by their study. For example, we now know that defect in pictorial interpretation can arise either from left occipital or right temporal lesions, but its basis appears quite different in the two cases. In the former, it appears to reflect a specific failure in multiple form perception; in the latter, a defect in intellectual analysis. It is evident, therefore, that much closer analysis of the causation of specific failures in visual perception and recognition is needed before one can argue with any propriety from the localization of defects to the localization of functions. None the less, Drs Hécaen and Angelergues have performed a very useful service in assembling this material and in drawing attention to the real and important differences between left-sided and right-sided cerebral lesions in the origin of higher visual disorders. Their work should be of great interest not only to clinical neurologists but also to psychologists interested in the contribution of clinical study to our understanding of brain function.

O. L. ZANGWILL

General Psychopathology. By KARL JASPERS. Translated from the German (seventh edition) by J. HOENIG and MARIAN W. HAMILTON. Manchester: University Press. 1963. Pp. xxxii + 922. 75*s*.

Gesammelte Schriften zur Psychopathologie. (Collected papers on psychopathology.) BY KARL JASPERS. Berlin, Göttingen and Heidelberg: Springer. 1963. Pp. vii + 420. DM 36.

Many philosophers have been interested in abnormal mental phenomena, but none of them has studied them as carefully and systematically and with so much clinical understanding as Karl Jaspers. He even took a medical degree and worked as an assistant physician in the Heidelberg psychiatric Klinik. He observed individual patients with immense industry and wrote case reports of unheard-of length and thoroughness. In 1913 he published his *Allgemeine Psychopathologie*, the 7th edition of which appeared in 1959. It has now been translated into English many years after it has made psychiatric history on the continent. The translation is superb. Psychopathology as defined by Jaspers is concerned with conscious psychic events. The study of the various ways in which they are experienced is called phenomenology, following Husserl. Psychic events can be understood 'from within', i.e. they can be meaningfully related to other psychic events. This is what Jaspers called 'the psychology of meaning', which deals with meaningful psychical connexions. In his opinion, psychoanalytic propositions belong to this category only. 'Explanatory psychology' is concerned with the causal connexions of psychic life. Causes are, according to Jaspers, always extra-conscious, which in this context means extra-psychic, i.e. organic. Abnormal mental phenomena can be attributed either to the 'development of a personality' or to 'a process', or to the combination of both. These two pairs of concepts are Jaspers' chief contributions to psychopathology, apart from his endeavour to analyse and clarify psychopathological phenomena, such as the various types of disorders of perceptions, delusions, dementia, etc. The book ranges widely over most of the fields which are of relevance to psychopathology. Although written 50 years ago, it is still worth reading. There are today few psychiatrists, in Germany or elsewhere, who would accept Jaspers' antitheses between the two types of psychology and between the two types of abnormal behaviour. Nevertheless, as a systematic presentation and as an attempt at a philosophical illumination of abnormal mental phenomena this work is unsurpassed. It is likely to remain unique as no single scholar could master so wide a range of subject matter today. Most probably the book was partly or wholly a reaction to the situation of European psychiatry around 1910. Freudian ideas had made a profound impact and had contributed substantially to Bleuler's concept of schizophrenia. On the other hand, the notion that 'mental diseases are brain diseases' still ruled supreme in most schools of psychiatry. Jaspers endeavoured to put everybody in his place, but in doing so he perpetuated existing notions about the etiology of mental disorders. His antitheses are as much out of date today as that of environmental versus hereditary factors. Like every great work, Jaspers' *Psychopathology* has illuminated some issues and obscured others. Being a philosopher, Jaspers was not concerned with progress but with critical appreciation and clarification of other people's ideas. He was opposed to all theories which failed to measure up to his standards. Kraepelin, Freud and Kretschmer were found wanting. Jaspers' existentialist orientation left little room for experimental psychology. He has had an inhibiting influence on German psychology and psychiatry. Unwittingly, he has delayed their emancipation from philosophy.

Gesammelte Schriften zur Psychopathologie is a series of eight papers published between 1909 and 1913. They contain the clinical observations from which Jaspers evolved his basic psychopathological concepts. They show Jaspers as a master of the phenomenologically orientated biographical case presentation. The collection of these classical papers is a valuable appendix to the General Psychopathology.

E. STENGEL

Clinical and Experimental Hypnosis. By W. S. KROGER. Philadelphia: Lippincott; Montreal: Pitman. 1963. Pp. xxii + 361. £5.

Hypnosis in Modern Medicine. Third edition, edited by J. M. SCHNECK. Springfield, Ill.: Thomas. 1963. Pp. xix + 452. $12.75.

The Nature of Hypnosis. (Transactions of the 1961 International Congress on Hypnosis.) Edited by M. V. KLINE. Baltimore: Waverley Press, Inc. 1962. Pp. x + 133. N.P.

Here are three newcomers to the serried ranks of books on hypnotism, most of which seem to say much the same things as Bramwell said over half a century ago. Of the three I should certainly recommend Kroger's *Clinical and Experimental Hypnosis* as the best value. His is a major textbook of wide coverage. The orientation is more towards the clinic than the laboratory. He recommends hypnotherapy for a huge variety of conditions and although, as in the case of epilepsy, it is sometimes glaringly obvious that his experience with certain types of patients is limited, his enthusiasm is tempered by common sense.

Hypnosis in Modern Medicine is written by a number of different authors whose contributions vary in quality. I thought the sections on history, dentistry and obstetrics sound and useful. The worst chapter, on hypnotherapy for children, was written by the only British contributor. The book is aimed at the physician, not the psychologist.

The Nature of Hypnosis is yet another of those collections of papers given at a conference where it was felt by those present that posterity would wish to read what they had said. For hypnotism fans only.

IAN OSWALD

Biological Prototypes and Synthetic Systems. Volume 1. Proceedings of the 2nd Annual Bionics Symposium, held at Cornell, 1961. Edited by EUGENE E. BERNARD and MORLEY R. KARE. New York: Plenum Press. 1962. Pp. xii + 397. $12.50.

This book is a collection of 48 papers, some of which are in abstract form only. The general subject has come to be called Bionics, the study of the relationship between living systems and synthetic mechanisms.

After an introduction by von Foester, there are five papers which report experimental findings. The only one of direct interest to psychologists is that by Tepas, Armington and Kropfl, which contains a description of a method for recording visually evoked potentials from the surface of the scalp. The rest of the book contains purely theoretical papers. Only those of interest to psychologists will be mentioned.

Most of the papers deal with perception mechanisms. Caldwell, Glaesser and Stewart describe the design of a model ear and Stewart outlines a theoretical model of hearing. Hildebrand describes an electronic model of the Limulus eye. All three papers are competent.

There are several papers on pattern discrimination, including papers by Sutro and by Vossler and Uhr which are discursive and trivial. Hu gives a mathematical model for shape constancy using 'moment invariants'. This is elegant but has no more theoretical value to a psychologist than less mathematical treatments. By contrast, Gibson discusses perceptual constancies descriptively but adds nothing to what he has written elsewhere. Blum describes a specific mechanism for pattern recognition which is, in essence, a scansion mechanism. Such models may enable the consequences of specific hypotheses to be worked out, but it is possible to make an infinite number of them to perform a given function, all based on different principles, without getting any closer to the understanding of actual biological processes. Pritchard describes how the stabilized retinal image technique has been used to study the mechanisms of figural organization. Olson and Belar discuss the recognition of speech by machine. Papers by Kanal and by Fu contain discussions of recognition in highly abstract terms.

There are a few papers on neurone networks. Mueller and Putzrath give a clear summary of this field. Two specialized papers by Yilmez and by Johnson are difficult to follow, possibly because of their style and jargon which includes new words such as 'percevonic' and 'neuromime'. Smith and Davidson discuss the implications of regarding neurone nets as randomly organized.

There are only three papers concerned with learning. One by Marzocco and Bartram gives a trivial analysis of a simple learning situation. Wiesen and Shuford report that subjects are able to estimate the proportion of items in a population when repeatedly presented with samples. This rather trivial finding is expressed in terms of a mathematical model (Bayesian strategies) which adds nothing to one's understanding. There is a competent paper on concept formation by Turner.

While many of the papers are stimulating and ingenious, several contain an over-elaborate jargon and use mathematical notation unnecessarily, thereby disguising the triviality of their actual content. But the exciting fact which this book demonstrates is that in the last few years there has been a rapid increase in the number of people in various disciplines whose interests have converged on the problems of living processes. In particular, the field of perception, which has been traditionally a branch of psychology, is more and more being studied by physicists, mathematicians and engineers and we can expect a rapid advance in our knowledge as a result of this cross fertilization of interests.

IAN P. HOWARD

Computer Simulation of Personality. Edited by S. S. TOMKINS and S. MESSICK. London: Wiley. 1963. Pp. x + 325. 40*s*.

This book contains the eight papers, and eleven short formal discussions thereof, presented to a conference held at Princeton in June 1962, and entitled 'Computer simulation: frontier of personality theory'. As might, perhaps, be expected, one of the papers (that of Uhr) is not concerned with personality at all, but with programmes which simulate the learning of simple pattern recognition and of the ability to translate, whilst another (by Blum) has nothing to do with computer simulation, and deals with experiments on subjects who have been extensively trained by hypnosis to experience specified degrees of anxiety, arousal and inhibition to order.

In the papers which are actually about the computer simulation of personality, quality appears to be highly negatively correlated with the authors' breadth of vision. Thus those of Colby and of Abelson, both on the simulation of the action of certain neurotic mechanisms on affect-laden beliefs, are narrow in scope, specific to the point of giving flow diagrams, and excellent. That of Loehlin, describing a model which responds to situations with numbers standing for the emotions attraction, anger and fear, and the corresponding behaviours approach, attack and withdrawal, is at a higher level of generality and much less satisfactory. That of Tomkins on the 'humanomaton', an R.U.R. type robot which has not been designed, or even clearly specified, but is rather an idea of what the optimal simulation might be, is vague and almost completely worthless. In addition to the papers on specific (or not so specific) models, there is one by Reitman on the more theoretical aspects of computer simulation. A resumé of the whole conference, by Messick, concludes the book.

J. P. N. PHILLIPS

Multivariate Procedures for The Behavioral Sciences. By W. W. COOLEY and P. R. LOHNES. London and New York: Wiley. 1963. Pp. x + 211. 51*s.*

This book will be of value to the psychologist who is interested in having his statistical data analysed by computer, and who desires to understand something of the process. It is intended to implement rather than constitute a course in multivariate analysis, and this it does quite successfully. The basic statistical methods of multiple regression analysis, canonical-correlation methods, analysis of variance, factor analysis, together with the underlying matrix algebra are covered concisely but clearly.

The main aim of the book is to help the reader to programme his material so that all the necessary statistical calculations associated with the foregoing techniques can be handled by I.B.M. computers. The programmes listed are written in the I.B.M. coding language known as 'Fortram'. As Hallworth says in a useful review ('Electric computer programmes for psychologists', *Bull. B.P.S.*, no. 48, 1962), the two languages of 'Fortram' and Mercury Autocode have gained by far the widest currency in this country, and although 'Fortram' programmes are not yet much used in university computers it is probable that they will be so used in the near future. If this is so it is clear that this book will be very useful in helping the research worker understand what he is about.

Flow diagrams and actual programmes are given in some detail for all the various techniques. Though the language and ideas will be rather strange to many, understanding will follow providing the reader is willing to work through the programme with actual data.

At the present time this book is of great value—one's only reservation is as to how long it will remain up to date in such a rapidly changing field.

JACK WRIGLEY

Introduction to Psychological Research. By WILLIAM A. SCOTT and MICHAEL WERTHEIMER. London and New York: Wiley. 1962. Pp. xiv + 445. 60*s.*

Methods of Psychiatric Research : An Introduction for Clinical Psychiatrists. Edited by P. SAINSBURY and N. KREITMAN. London: Oxford University Press. 1963· Pp. xii + 175. 35*s.*

Basic Ideas of Scientific Sampling. By ALAN STUART. London: Griffins. 1962. (Griffin's Statistical Monographs and Courses, No. 4.) Pp. 99. 18*s.*

The Method of Paired Comparisons. By H. A. DAVID. London: Griffin. 1963. (Same series, No. 12.) Pp. 124. 28*s.*

Tables of Testing Significance in a 2 × 2 Contingency Table. Compiled by E. S. PEARSON. Cambridge: University Press. 1963. Pp. vi + 103. 16*s.*

It is seldom that one encounters two books so clearly similar in scope and intention, yet different in provenance, authorship and incidental detail as the two that head the list of titles given above. The similarities arise from their common origin in the contemporary need for expansion not only of higher education in the general sense but of facilities for research and of the number of people who engage in it. Closely similar, too, are the broad outlines of the structure of each book: two or three introductory chapters on what research is and how it is administered: the bulk of the book, on the design of experiments and methods for the analysis of data; a concluding chapter on the preparation of a scientific paper. Even the number of chapters—fourteen —is, no doubt by coincidence, the same.

Within this pattern the variations are more or less as expected. Scott and Wertheimer pay more attention to details of statistical treatment; most of the ground covered is familiar, but with more emphasis than usual on underlying principles. Perhaps the fact that the book is the product of a true collaboration has given it a unity which is somewhat lacking in *Methods of Psychiatric Research,* in which each chapter is written by a different author, and in which the weighty tends to alternate with the relatively trivial. Some (although by no means all) of the subject matter is also more specialized, as was to be expected.

Both books are thoroughly recommendable: if one must choose, Scott and Wertheimer's is to be preferred, not so much because it deals with matters 'psychological' rather than 'psychiatric', but for the reasons hinted at above, and because generally speaking it 'gives better value'. Where Sainsbury and Kreitman score is in their invaluable account of British sources of financial support, etc. Scott and Wertheimer, incidentally, do the same for the U.S.; each book is (inevitably?) rather firmly rooted on its own side of the Atlantic.

The other works listed are likely to be of great value to research workers in psychology or in any of the biological or social sciences. Stuart's book presents a particularly lucid minimally mathematical account of the principles of sampling; these, as the author says, are 'illustratively expounded rather than demonstrated'. David's work is frankly more statistical: 'Writing as a statistician I have avoided technical terms *from other fields*' (reviewer's italics). The language is definitely that of the mathematician. Nevertheless, Professor David makes ample reference to practical applications of the method.

The purpose of the *Tables* listed is clear from their title. Similar contingency tables for small samples have been readily available: the present work gives full data for groups of $N = A+B$ individuals where $B \leqslant A \leqslant 30$, and abridged data up to $31 \leqslant A \leqslant 40, B \leqslant A$. The typography is well up to the high standard one expects of the Cambridge University Press, and the work is very reasonably priced.

B. SEMEONOFF

Appraising Vocational Fitness by Means of Psychological Tests. New York and London: Harper and Row. Revised edition, 1963. Pp. xvi + 688. 96*s*.

This book consists of two volumes rolled into one. The first volume discusses the nature of intelligence, of aptitudes, of interests, and of personality. It gives sufficient indication of how tests are, or ought to be, constructed, validated, and used, to enable the reader to judge for himself how satisfactory a particular test should be for vocational guidance, or for selection purposes. And it presents a systematic approach to vocational counselling. This volume is written clearly, concisely and with discrimination.

The second volume consists of a mass of detailed information (chunks of which are incorporated in each of the middle chapters), about certain tests which are neither comprehensive nor carefully selected. For example, Chapter VIII on 'Clerical Aptitude' tests is confined almost entirely to the Minnesota Clerical test, which is neither unique nor the best of its kind.

The book surveys a variety of tests, both cognitive and of temperament, and contrasts general aptitude test batteries with *ad hoc* collections of tests for twelve specific occupations. In the key chapters, XX–XXII, the authors describe their methods of vocational appraisal, how to present and interpret test results in counselling, and the preparation of reports to professional workers and clients. Finally, Appendices A and B contain four specimen case histories. The original appraisals in Appendix A were written before the follow-up data in Appendix B became available, so as to avoid contamination by hindsight. This material therefore, in substance and in presentation, would be especially valuable to anyone with the opportunity and desire to record his clients' progress with a view to increasing his own efficiency. There is a comprehensive bibliography of more than a thousand references, and separate indexes for Test Publishers, Authors, Occupations, Subjects, and Tests.

Throughout the emphasis is on written tests and on situations where many test results are already available or there is ample time for further tests to be given. Interviewing receives inadequate treatment, and the 'extended interview' procedure is mentioned hardly at all. Conspicuous by their absence are any references to the encouraging results in some British follow-up studies, e.g. the over-all validities of 0·5 to 0·6 given by P. E. Vernon in 'The Validation of Civil Service Selection Board Procedures', *Occupational Psychology*, **20**, 1950. Nor is attention drawn to the difficulties of continually inventing fresh test items—one imaginative test constructor is worth more than several scoring machines! Despite this lack of balance, the book would be extremely valuable to a vocational counsellor, but to anyone else it would be useful only for occasional reference purposes.

E. ANSTEY

The Quick Test (QT): Provisional Manual. By R. B. Ammons and G. M. Ammons. *Psychological Reports Monograph Supplement*, 1–V 11. 1962. Pp. 53. $2.50.

The *Quick Test* is essentially a shortened version of the same author's *Full Range Picture Vocabulary Test* (*FRPV*), reviewed in the *Fourth Mental Measurements Yearbook* (but no reference in the *Fifth*). Each of its three 'forms' is represented by a card bearing four line drawings. The subject is required to say, for each of 50 words associated with a given card, which of the four pictures 'best fits it'. The purpose of the test is to provide a means of 'quick screening of verbal intelligence in practical situations'. Separate norms (in the form of 'I.Q.' conversions) are provided for each form separately, any pair, or all three together.

The idea is an attractive one, but the material published is open to criticism. Three points, in order of increasing seriousness, are as follows: First, the drawings are crude—though seldom really ambiguous. Secondly, some of the attributions of meaning, and the cues on which they rest, are highly dubious. Thus, *celerity* is regarded as appropriate to a picture of a restaurant because the waitress is 'bending over in her hurry'. Again, in the same picture the diner very definitely looks *bovine*—another of the key words, for which the correct choice is a drawing showing a cow by a stream, because, the manual says, this 'is the only drawing with a cow in it'. Thirdly, and finally, the method seems positively to encourage acceptance of loose use of language. Whether this is inherent in the method would seem to rest on further experiment with better pictures. Here everything seems to have been sacrificed to convenience, and one cannot help feeling that the validities claimed for the test and its parent FRPV (which are considerable) must have been achieved in spite of rather than because of their characteristic features.

Other recent monographs (1962–3) in this series (see review in this *Journal*, 1962, **53**, 355–6) of more than routine interest include:

Psychological Reports: Genes, Drives and Intellect, by K. J. Hayes (2–V 10); *Parsimony in Psychology*, by W. F. Battig (3–V 11); *Persuasion and the Acquisition of Attitudes: Models from conditioning and selective Learning*, by R. F. Weiss (4–V 11); *Aversive Conditioning Methodology in Animal Research*, by M. L. Goldstein (5–V 11).

Perceptual and Motor Skills: Factor Analysis and the Investigation of Hypotheses, by H. P. Bechtoldt (2–V 14); *Odor and Olfaction: a Bibliography, 1948–60*, by K. M. Michels and others (6–V 15). *Studies in Intercultural Testing*, by S. D. Porteus and A. J. Gregor (7–V 16).

Most of the above, incidentally, contain extensive bibliographies. They are priced at from $1.00 to $3.00.

B. SEMEONOFF

Pickford Projective Pictures. By R. W. Pickford. London: Tavistock. 1963. Pp. xi + 122. Text 30*s*. Pictures 25*s*. The Set, 50*s*.

The Pickford Projective Pictures are a collection of 120 post-card size reproductions of charcoal line-drawings showing children in a wide variety of situations involving other children, parents, physical dangers, fantastic animals, etc. They are intended for use with pre-adolescent children who are asked 'to tell a story as long and imaginative as possible' about each of the six cards offered at one session. The stories—which, it is said, are to be memorized by the psychologist rather than written down in the child's presence—may then be used for diagnostic purposes. In addition, they may—and this is the author's more interesting claim—be used as a vehicle for psychotherapy. The continuous session-by-session story-telling process will be helpful to Rogerian and Analyst alike and can be helpful also in the treatment of quite specific educational difficulties.

The accompanying test describes the administration of some of the pictures to 85 boys and 44 girls referred to Glasgow Child Guidance Clinics with such complaints as enuresis and speech disorders. Normal children's responses are not described. A catalogue is given of the 4 or 5 major response types found for each of the 120 pictures but no attempt is made to relate response-type or frequency with diagnosis. Indeed, the author's major concern is with the therapeutic use of the material and a careful account is given of judgements of the material'usefulness by psychologists who have used it in various contexts. There are five detailed cases histories illustrating the use of the pictures in therapy.

JOHN LIGGETT

British Journal of Industrial Relations. Volume 1, No. 1, Feb. 1963; No. 2, June 1963. London: London School of Economics and Political Science. 1963. Pp. 291. 15*s*. a copy, 42*s*. one year.

Scholarly, up-to-date, readable: these describe this new Journal which fills a gap between (say) the *Journal of the Institute of Personnel Management* and the 'heavier' journals serving Psychology, Sociology, and Economics.

The London School of Economics and Political Science—the members of its staff are the major contributors—is to be congratulated on publishing this Journal which, although not of immediate interest to all psychologists, brings to the attention of those working in the applied field the problems which would repay their specialized research attention. It is obligatory reading, or scanning, for psychologists working in inter-disciplinary teams and institutions.

DENIS MCMAHON

Psychoevaluation: Adaptation—Distribution—Adjustment. By MILTON E. HAHN. New York and London: McGraw-Hill. 1963. Pp. x+228. 62*s*.

If one of Maslow's self-actualizing persons decided that for some reason he required the services of a professional psychologist to whom could he turn? Professor Hahn describes at length the inadequacies of present American counselling and psychotherapeutic practice for such a person. The substance of his book is the plea for a new technique which he would call psychoevaluation, and for new training facilities for this technique, with, in addition, adequate official recognition for the new discipline. The technical mysteries of psychoevaluation he reserves for a forthcoming publication. In this country we have recently received a salutary lesson in democracy; mere title, wealth and noble birth must not debar a man from high political office. Accordingly our profession must see to it that we do not deny our services to the mature, successful, normal, self-actualizing men and women amongst us. However, until Professor Hahn's next work appears the nature of these services must remain uncertain. The present volume is perhaps best regarded as an extended advertiser's announcement.

J. DREWERY

Committees: How They Work and How to Work Them. By EDGAR ANSTEY. London: Allen and Unwin. 1962. Pp. 116. 15*s*.

This is an excellent little book. It is amusingly illustrated by Thelwell, and while these drawings are entertaining in themselves they may be slightly misleading. This is not basically a humorous work but a serious attempt to provide a short, practical guide for the committee member, secretary and chairman.

Dr Anstey examines the structure and functions of committees but is largely concerned to provide a useful account of the group dynamics of the decision-taking process. He does not describe his book in this way and one suspects that he would be loth to do so for he is studiously non-technical and non-theoretical and his book gains in clarity from its essentially practical exposition. This book should be on the agenda of the next committee you attend.

J. DREWERY

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review.)

The Psychology of Crime. By DAVID ABRAHAMSEN. New York: Columbia University Press. London: Oxford University Press. 1960. Pp. xiv + 358. 48*s.*

Understanding Organizational Behaviour. By CHRIS ARGYRIS. London: Tavistock. 1960. Pp. xii + 179. 38*s.*

Metaphor and Symbol. Proceedings of the 12th Symposium of the Colston Research Society held in the University of Bristol, 28–31 March 1960. Edited by L. C. KNIGHTS and BASIL COTTLE. London: Butterworths Scientific Publications. 1960. Pp. xii + 150. 30*s.*

Origins of Alcoholism. By WM. McCORD and JOAN McCORD. London: Tavistock. 1960. Pp. xiv + 193. 25*s.*

Handbook of Research Methods in Child Development. By PAUL H. MUSSEN. New York: Wiley. 1960. Pp. x + 1061. 122*s.*

Clinical Child Psychiatry. By KENNETH SODDY. London: Baillière, Tindall and Cox. 1960. Pp. x + 470. 42*s.*

The Healthy Child. Edited by HAROLD C. STUART and DANE G. PRUGH. Havard: University Press. London: Oxford University Press. 1960. Pp. xvi + 507. 80*s.*

Delinquency and Opportunity. By R. A. CLOWARD and L. E. OHLIN. London: Routledge and Kegan Paul. 1961. Pp. xii + 220. 25*s.*

Subnormal Personalities. By C. J. C. EARL. London: Baillière, Tindall and Cox. 1961. Pp. xiv + 338. 30*s.*

Imagination. By E. J. FURLONG. London: George Allen and Unwin. 1961. Pp. 125. 18*s.*

The Analysis of Behavior. By JAMES G. HOLLAND and B. F. SKINNER. New York and London: McGraw-Hill. 1961. Pp. xvi + 175. 27*s.*

Man's Picture of His World. By R. E. MONEY-KYRLE. London: Duckworth. 1961. Pp. 190. 21*s.*

Childhood in Autobiography. By EDNA OAKESHOTT. Cambridge: The National Book League—Reader's Guides—Fourth Series No. 1. 191. Pp. 24. 3*s.* 6*d.*

Leadership and Organization. By ROBERT TANNENBAUM, IRVING R. WESCHLER and FRED MASSARIK. New York and London: McGraw-Hill. 1961. Pp. xiv + 456. 58*s.*

On Knowing: Essays for the Left Hand. By JEROME S. BRUNER. Cambridge, Mass.: Harvard University Press. 1962. Pp. 165. 30*s.*

Text and Cases in Marketing. By EDWARD C. BURSK. London: Prentice-Hall International. 1963. Pp. xii + 580. 42*s.*

The Origins of Scientific Sociology. By JOHN MADGE. London: Tavistock. 1963. Pp. 600. 63*s.*

The Enterprise and its Environment. By A. K. RICE. London: Tavistock. 1963. Pp. xiv + 364. 63*s.*

Managers. Personality and Performance. By KENN ROGERS. London: Tavistock. 1963. Pp. vii + 184. 30*s.*

Organizational Choice. Capabilities of Groups at the Coal Face under Changing Technologies. By E. L. TRIST *et al.* London: Tavistock. 1963. Pp. xv + 332.

The Mental Patient Comes Home. By HOWARD E. FREEMAN and OZZIE G. SIMMONS. London: Wiley. 1963. Pp. xii + 309. 60*s.*

Bio-Psychology: a Biological Approach to Human Motivation. New York: Exposition Press. 1963. Pp. 216. $6.00.

Over the Sex Border. By GEORGINA TURTLE. London: Gollancz. 1963. Pp. 319. 30*s*.

On Fundamental Measurement in Psychology. By GUNNAR GOUDE. Stockholm: Almqvist and Wiksell. 1962. Pp. 176. Sw.Kr. 18.

Elemente Einer Akustischen Theorie Der Vokalartikulation. By GEROLD UNGEHEUER. Berlin: Springer. 1962. Pp. viii + 124. DM. 24.80.

Konflict und Konfliktschicksal. By PETER BRUCHNER. Bern and Stuttgart: Huber. 1963. Pp. 120. S.Fr. 19.

Rules, Perception and Intelligibility. By F. A. HAYEK. London: Oxford University Press. 1963. Pp. 322-44. 5*s*.

Teaching Machines and their Use in Industry. By H. KAY, J. ANNETT and M. E. SIME. London: Her Majesty's Stationery Office. 1963. Pp. 30. 3*s*.

The Interprofessional Treatment of the Disturbed and Delinquent Adolescent. (Annual of the New York Academy of Sciences, Vol. 105, Art. 6.) New York: New York Academy of Sciences. 1963. Pp. 386–419. $2.50.

Formal Models for a Central Prediction System. (Psychometric Monographs No. 10.) By LEDYARD R. TUCKER. Richmond, Va: William Byrd Press. 1963. Pp. xi + 61. $4.00.

Development of Managerial Performance. A Research Approach. (Monograph C-9, 1963.) By THOMAS MAHONEY, THOMAS JERDEE and STEPHEN CARROLL. Cincinnati: South-Western Publishing Co. 1963. Pp. viii + 67. N.P.

Normalisierung Einer Projektiven Prüfungsmethode. Gruppen- und Individualversuche mit dem Helsingforstest. By KARL BRUHN. Helsinki: Societas Scientiarum Fennica. 1962. Pp. 242. Mk 1210.

Intelligenzprüfungen an Menschenaffen. 2. Auflage. By WOLFGANG KOHLER. Berlin: Springer. 1963. Pp. 234. DM. 28.

Conceptele Figurale. By E. FISCHBEIN. Bucharest: Editura Academiei Republicii Populare Romine. 1963. Pp. 473. Lei 16, 30.

Tratat de Psihologie Experimentala. By ALEXANDRU ROSCA (ed.). Bucharest: Editura Academiei Republicii Populare Romine. 1963. Pp. 650 + abstract in English. Pp. 30. Lei 35, 40.

Schmerz und Schmerzhaftigkeit. By ALFRED PRINZ AUERSPERG. Berlin: Springer. 1963. Pp. vii + 76. DM. 12.80.

Auf dem Wege zu einer Anthropologischen Psychiatrie. By JURG ZUTT. Berlin: Springer. 1963. Pp. viii + 570. DM. 48.

Eignungs- und Erziehungsdiagnosen für höhere Schulen mit dem Diapostiv-Z Test. By ROBERT BURGER. Stuttgart: Ernst Klett. 1963. Pp. 193.

Our Adolescents: their Interests and Education. By R. S. PANDEY. Agra: Agarwal. 1963. Pp. xviii + 387. Rs. 15.

Psychology. The Study of Behavior. By PAUL SWARTZ. Princeton, New York and London: Van Nostrand. 1963. Pp. xix + 451. 54*s*. 6*d*.

Patterns of Psychopathology. By MELVIN ZAX and GEORGE STRICKER. New York: The Macmillan Company; London: Collier-Macmillan. 1963. Pp. iv + 314. 25*s*.

Textbooks of Psychiatry for Medical Practice. By C. F. HOFLING. London: Pitman Medical Publishing Co. Ltd. 1963. Pp. xv + 558. 75*s*.

The Dogma of Christ. By ERICH FROMM. London: Routledge and Kegan Paul. 1963. Pp. viii + 151. 18*s*.

Salesmanship: Concepts, Management, and Strategy. By WILLIARD M. THOMPSON. London: Wiley. 1963. Pp. x + 530. 60*s*.

Vol. 55. Part 2 May 1964

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
BORIS SEMEONOFF

AIDED IN THE SELECTION OF PAPERS BY
C. BURT
AND

F. C. BARTLETT
E. A. BOTT
D. E. BROADBENT
R. S. CREED
J. DREVER

G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER

T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
M. D. VERNON

ASSISTANT EDITOR
R. M. McKENZIE

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W. 1
AMERICAN BRANCH, 32 EAST 57TH STREET, NEW YORK 22, N.Y.

Price Twenty-five Shillings net (U.S.A. $4.25)

Brit. J. Psychol.

Preparation of Papers for the British Journal of Psychology

The *British Journal of Psychology* subscribes to the *Fair Copying Declaration* of the Royal Society under the terms of which reproduction of material published in this Journal is permitted.

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. All papers, including reviews, must be typed (double spacing) on one side of the paper only and the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of most articles published from Vol. 47 (1956) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

2. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs etc. is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered caption of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following example:

Galton, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.
Révész, G. (1951). Le problème du génie. *L'Année psychol.* 50, 83–96.

Note that World List abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

Brit. J. Psychol. (1964), 55, 2, *pp.* 121–132
Printed in Great Britain

CALCUTTA UNIV. CENTRAL LIBRARY

RESPONSE INSTRUCTIONS AND THE SPEED OF RELATIVE JUDGEMENTS

II. PITCH DISCRIMINATION*

BY C. P. WALLIS AND R. J. AUDLEY
University College London

Subjects were required to make easy pitch discriminations at four frequency levels. In a counterbalanced design, for half their judgements they were required to choose the higher of the pair of tones, and for the other half, the lower. It was found that subjects were quicker at choosing the higher than the lower of two high tones, but quicker at choosing the lower than the higher of two low tones. This 'cross-over' effect was found to be graded across the stimulus levels intermediate in the series, the intersection point being somewhat below the geometric mean of the series of frequencies used. While latencies tended to decrease during the course of an individual session, the cross-over effect was consistently observable in an analysis that controlled for such practice effects. Whether the tone responded to was the first or the second of the pair discriminated was found to be a significant variable. The 'cross-over' phenomenon was only consistently effective when the second tone was chosen. Linking these findings to similar phenomena previously demonstrated in brightness discrimination, theoretical suggestions are advanced which are relevant to all judgemental processes.

When subjects are required to make a brightness discrimination, their judgement time is influenced by an interaction between the response category demanded of them and the relative level of intensity of the pair of stimuli judged (Audley & Wallis, 1964). When the pair of stimuli to be discriminated is relatively intense subjects are quicker at choosing the brighter than they are at choosing the darker, but the choice of the darker is quicker when the stimuli are relatively dark. It was suggested that this kind of phenomenon, which was referred to as a '*cross-over*' effect, is probably common to many situations involving judgement.

This suggestion would be strongly endorsed if a similar effect could be shown to operate in a different sense modality. The pitch dimension, which requires a relatively simple judgement and involves an easily manipulable stimulus, was therefore selected for investigation. There are a number of important ways in which judgement of pitch and brightness differ. First, unlike the 'prothetic' and summative dimension of brightness, the pitch dimension is 'metathetic', variations along the subjective scale being presumed to be in some way 'qualitative' in nature. It has been suggested (Stevens, 1957) that quite different types of judgemental system might underly the two sorts of dimension. The time order errors, for example, which are normally found in comparative judgements made with 'prothetic' stimuli (e.g. Leuba, 1892, for brightness), have been denied for judgements of pitch (Koester, 1945; Postman, 1946) though again such errors have been claimed in other studies (Whipple, 1901; Wada, 1932; Koester & Schoenfeld, 1946). Secondly, while in our previous experiments on brightness judgements were of simultaneously presented stimuli, in pitch discrimination the pair of tones judged are necessarily separated in time. The design of the present auditory experiment has controlled against the

---

* The research reported in this document has been sponsored by, or in part by, the 'Air Force Office of Scientific Research, OAR' through the European Office Aerospace Research, United States Air Force.

influence of classical time order errors, but has allowed for the analysis of the effect of temporal ordering of the stimuli upon the processes of pitch discrimination. Finally, the luminance of the background to the stimuli in the brightness experiment was found to play an important role. In the case of tones presented for pitch discrimination it is doubtful whether any physically relevant background exists. Clearly, if the 'cross-over' effect could be shown to operate in judgements of pitch differences, this would provide strong evidence for the generality of the phenomenon.

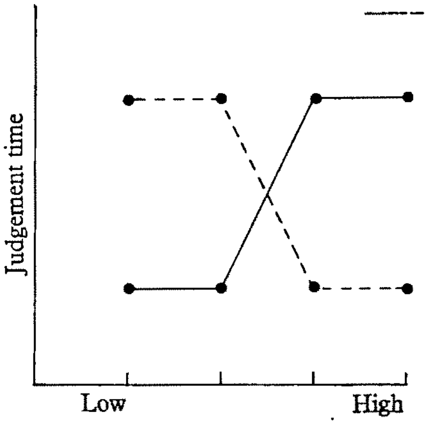


Fig. 1

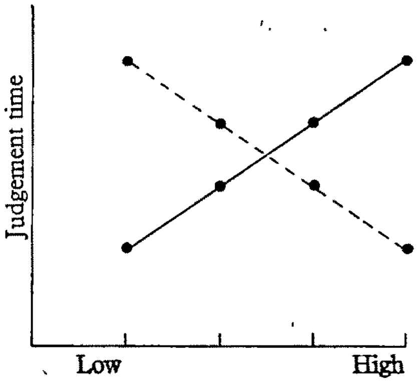


Fig. 2

In most of the studies previously reported only two brightness levels were used. In the auditory experiment a greater range of frequency variations has been used, enabling an assessment to be made of the extent to which the 'cross-over' phenomenon, the difference in the judgement times between choosing the 'higher' and the 'lower' tone, is operative for stimuli in the middle of the stimulus series.

Three possible effects might be expected over a range of different discriminations:

(1) The disparity in judgement times for the two kinds of response might be found only for extreme stimuli.

(2) The intermediate stimulus levels might show a cross-over disparity in the same direction and to the same extent as the end stimulus nearest to which the intermediate stimulus lies. This is shown in Fig. 1.

(3) The cross-over disparity could be graded over the intermediate levels, which show the effect in the same direction as, but to a lesser extent than, the nearer end stimulus. This is illustrated in Fig. 2.

Two experiments on pitch discrimination will be described, a pilot study (AI) using 5 frequency levels, and the main experiment (AII) using 4 levels. In this report the pitch levels will be numbered from 'low' to 'high', the numeral 1 being allotted to the lowest pitch, and 4 or 5 to the highest.

EXPERIMENT AI (PILOT)

Method

Design. At each of 5 pitch levels subjects were presented with 12 pairs of tones at 60 db. intensity providing an easy pitch discrimination. For half the presentations the subject was instructed to choose the higher tone, and for half the lower, subjects being split into 2 counterbalanced groups on the basis of the order of the response category. The HLLH group made 15 responses of 'which higher', 30 of 'which lower' and a final 15 of 'which higher'; for the LHHL groups the order was reversed, starting and finishing with 15 responses of 'which lower'.

The following factors were randomized within each block of responses;

(1) The first tone being higher or lower than the second.

(2) The occurrence of a pair of tones at one of the 5 frequency levels used.

On half the presentations the subject was required to respond to the first tone of the judged pair, and on the other half to the second tone. This controlled against classical time order effects, which should also have been reduced by the short time interval of a half second between the two stimuli.

Apparatus

Three Nagard time switches controlled two Peter's audiometers to provide two tones each of 0·5 sec. duration with a 0·5 sec. interval between them. Each tone could be manipulated by the experimenter separately, and was presented binaurally to the subject through earphones. The time switch that controlled the onset of the second tone simultaneously started a chronotron, which could be stopped by the subject depressing either of two response keys. In the event of a quick reaction time, the chronotron could be stopped before the termination of the second tone.

The five frequency levels used (125, 750, 2000, 4000 and 7000 c.p.s.) were chosen to cover the range of audible frequencies at approximately equal subjective intervals. At each level a pitch interval between the pair of stimuli was selected to provide an easily discriminable difference that was of approximately equal difficulty at all levels, and was maintained objectively the same for all subjects. The frequency differences used, running from the low to high frequency levels, were +12·5, +80, +350, +900, and +1500 c.p.s. These differences were added to the basic frequency level.

Procedure

The subject wore headphones through which he heard binaurally a pair of tones which differed in pitch. Each stimulus presentation was preceded by a warning light. Left and right response keys were provided. When making a judgement of which tone was higher the subject was instructed to press the right key if the first tone was higher in pitch and to press the left key if the second tone was higher. Similarly for choice of the lower tone, the right key was pressed when the first tone was lower and the left key when the lower tone came second. As an aid to memory these instructions remained written on a board which was in front of the subject throughout each session.

The subject was instructed to make his judgement as quickly as possible and was informed that his reaction time from the onset of the second tone would be measured. 60 response times were recorded for each subject.

Subjects

Six male and 6 female subjects were used, with an even distribution of both sexes in the two counterbalanced groups. The subjects were university undergraduates.

Problems with the Pilot Study

The subjects found some difficulty in remembering the instructions and in conforming to them. Essentially a difference in time (first and second tone) had to be translated into a difference in space (left or right key) and at the same time the response 'set' of choosing either the higher or the lower had to be maintained. Introspective reports showed that many subjects had recourse to visual imagery to assist them in their task, picturing the keys, for example, as parts of a piano key-board. As a key-board simultaneously contains both a 'higher' and a 'lower' element, use of this type of mnemonic partly defeated the attempt to manipulate the response variable.

Another difficulty occurred with the highest pair of stimuli (7000+1500 c.p.s.) which was too close to the upper pitch threshold for some subjects. A greater variation in judgement time was found for this discrimination, and the mean response latency was approximately 100 m./sec. longer than latencies for the other pitch levels.

Results

For the condition where the higher tone was chosen, mean latencies can be calculated for each of the 5 pitch levels, and these can be compared with latencies for the choice of the lower tone for the same discriminations. The hypothesis under examination predicts a 'cross-over' of these latency curves, the choice of the higher tone being quicker for the discrimination of high frequencies and the choice of the lower being quicker for low frequencies. The results are exhibited in Table 1 and Fig. 3.

Table 1. *Experiment A I. Mean judgement times at 5 frequency levels under two response instructions in seconds*

	Frequency level				
Response category	I	II	III	IV	V
Which lower	0·645	0·626	0·644	0·788	0·881
Which higher	0·843	0·717	0·752	0·787	0·893

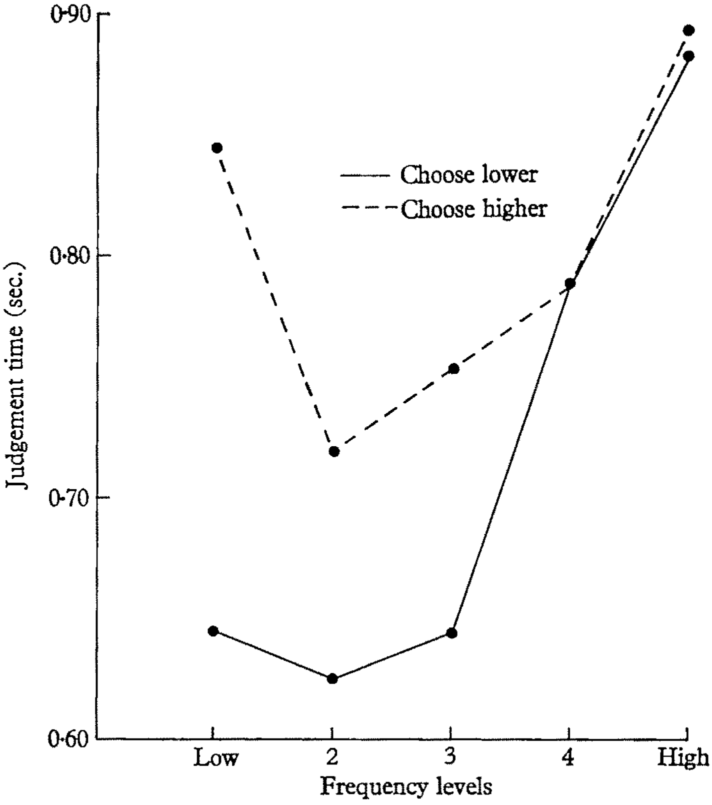


Fig. 3. Pilot. Mean judgement times of 5 tone pairs presented binaurally.

The 5 frequency levels have been placed at equal intervals along the abscissa as these frequencies are approximately equidistant along the mel scale of subjective pitch (Stevens & Volkmann, 1940).

Despite the difficulties mentioned above there is a marked disparity in response latencies at the low pitch levels, an average response of 'which lower' being 23·5 % quicker than that for the choice of 'which higher' for the lowest pair. At levels II and III the response of 'lower' is quicker by 12·7 and 14·4 %, respectively, while differences at the two highest levels are minimal. This 'funnel effect' is similar to that found in the earlier studies of brightness discrimination, particularly Expt. I of the previous paper, in which three intensity levels were used, all brighter than the background. However, in view of the subjects' difficulty with the task a modification of the experimental procedure was obviously required.

EXPERIMENT A II

Design

The design and apparatus of the main experiment on pitch discrimination were the same as those used in the pilot, except for the following changes:

(1) The highest pitch level (7000 c.p.s.) was discarded, the other 4 frequency levels and discrimination differences being kept unchanged.

(2) The tones were delivered through the headphones monaurally, the two individual tones of any discrimination pair being delivered in succession to alternate ears. The tone durations and the interval between the members of a pair were kept at 0·5 sec.

(3) Within each response instruction block of presentations in the counterbalanced design (HLLH and LHHL) the following factors were randomized; (*a*) the first tone being delivered to the left or right ear: (*b*) the first tone being higher or lower than the second: (*c*) the occurrence of a pair of tones at one of the 4 frequency levels.

(4) Each subject received 64 presentations, 8 judgements being made at each of the four pitch levels under each of the two response instruction conditions.

Procedure

Following a warning light the subject received through the headphones a pair of successive tones differing in pitch. One tone was delivered to the left ear and the other to the right, which ear received the first tone being randomized. The subject was required either to judge which of the pair was higher or to judge which was lower in pitch, the response instruction varying in accordance with the counterbalanced design. He was required to respond as quickly as possible by pressing that member of a pair of response keys which was on the same side as the appropriate higher or lower tone. The response latency, measured from the onset of the second stimulus, was recorded.

Subjects

Another group of 6 male and 6 female undergraduate students were used, distributed evenly by sex in the two counterbalanced groups.

Results

For each pitch level the judgement times under the two response conditions are comparable in Table 2 and Fig. 4.

Under the new conditions the full 'cross-over' effect was obtained, judgements of 'which lower' being quicker for the two lower discriminations, and judgements of 'which higher' being quicker for the two higher discriminations. To assess the extent of this difference, judgement times for the two extreme pitch levels were compared. Under the response instruction of 'lower' for each subject the mean judgement time for stimulus IV was subtracted from his mean for stimulus I, and a similar difference of the means calculated for a response of 'higher'. The difference between these differences was then calculated, and the sign of this difference was found to be in the direction of our hypothesis for 11 out of the 12 subjects. A test for correlated means gave a t of 3·7, significant at the 0·005 level with a 1 tail prediction. The subjects were therefore relatively quicker at choosing the higher of an extremely high pair of tones, and relatively quicker at choosing the lower of an extremely low pair.

It should be borne in mind that the slopes of the lines representing the times of the two kinds of judgements in Fig. 4 are partly determined by the relative difficulty of the four discriminations used. The frequency difference presented at one level could, for a given subject, be more easily detected than the difference at another frequency level, and for different subjects there would be variations in the relative ease of the four discriminations. The movement of one curve relative to the other, however, and the point of intersection should be determined only by the 'cross-over' effect.

Table 2. *Experiment II. Mean judgement times in seconds at 4 frequency levels under two response fluctuations*

Response category	I	II	III	IV
Which lower	0·608	0·612	0·650	0·716
Which higher	0·733	0·654	0·566	0·635

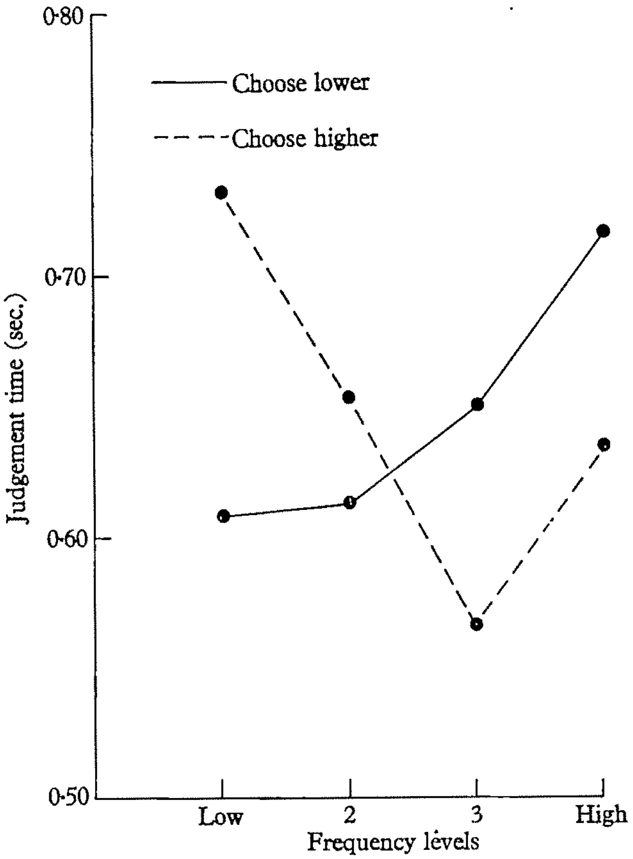


Fig. 4. Expt. AII. Mean judgement times of 4 tone pairs with comparison pairs presented to alternate ears.

Series effects

An examination of the overall means (Table 2) shows that the cross-over point lies between the second and third pitch levels. A t test for correlated means was applied to the mean response times for these two middle stimuli under the two instruction conditions. Again the signs of the difference between the differences was in the direction of the hypothesis for all but 1 of the 12 subjects. A *t* of 3·12 was obtained, giving a 1 tail significance at the 0·003 level. At the two intermediate frequency levels, then, subjects are also quicker at choosing the higher of a relatively high pair of tones, and the lower of a relatively low pair.

Of the three expectations considered in the introduction, the first can be dismissed. Stimuli lying within a series are affected by the 'cross-over' phenomenon. An immediate examination of the means, however, does not enable a discrimination to be made between the second and third suggestions. In moving up the scale from the lowest pitch level the effect appears to be graded. Thus a response of 'lower' to stimuli I was on the average 17·1 % quicker than a response of 'higher', while to stimuli II a response of 'lower' was only 6·9 % quicker. Moving down from the fourth pitch level, however, an all-or-none effect is suggested. A response of 'higher' was quicker by 11·3 % for stimuli IV, and by 12·9 % for the lower stimuli III. Further support for the all-or-none suggestion is furnished by the high level of significance found on comparing the 'difference between the differences' for pitch levels II and III, a significance level as high as that found on comparing pitch I with pitch IV.

To assess the extent to which the effect is graded, for each subject the difference in mean latency under the two response conditions was calculated for the 4 pitch levels. The four differences found for each subject were then placed in size order: according to the 'graded' hypothesis, the rank order of the differences should coincide with the stimulus order. A 'tau' coefficient was calculated for each subject, a tau of +1·0 showing a perfectly consistent rank order across the stimuli levels, and a tau of zero showing complete inconsistency. Seven subjects were found to have a tau of +1, 1 subject had a tau of 0·67, 3 of 0·33 and 1 subject a tau of zero. No subject had a negative tau. This examination of the change of the individual means across the stimulus scale provides strong support for a graded effect. The theoretical significance of this will be discussed later.

Practice effects

During the course of an experimental session a subject's reaction time tends to diminish. To prevent this having any systematic effect on the results two counterbalanced groups were used, an HLLH group starting and finishing with blocks of 16 responses of 'which higher' and an LHHL group starting with 16 responses of 'which lower'. Figs. 5 and 6 show, for each counterbalanced group, the mean judgement times across the 4 pitch levels for each response block, these blocks being arranged along the abscissa in time order.

In both figures, the general trend over the four curves from top left to bottom right shows the overall speeding up in judgement times. Consider for the HLLH group (Fig. 5) the latencies for the extreme pitch levels. With their first block of 'which higher' responses subjects are relatively slow, but are quicker in discriminating at the highest pitch level than at the lowest. As they moved through the middle blocks of 'which lower' responses, reaction times decreased, but much more markedly for the lowest stimulus. For the final block of H responses this relative movement is again reversed, level IV once more proving the quickest discrimination. It can thus be seen that, despite practice effects, the 'cross-over' phenomenon is operating uniformly in each block of responses. Similar results are found for the intermediate frequency levels in the HLLH graph, and the four frequency levels in the LHHL graph.

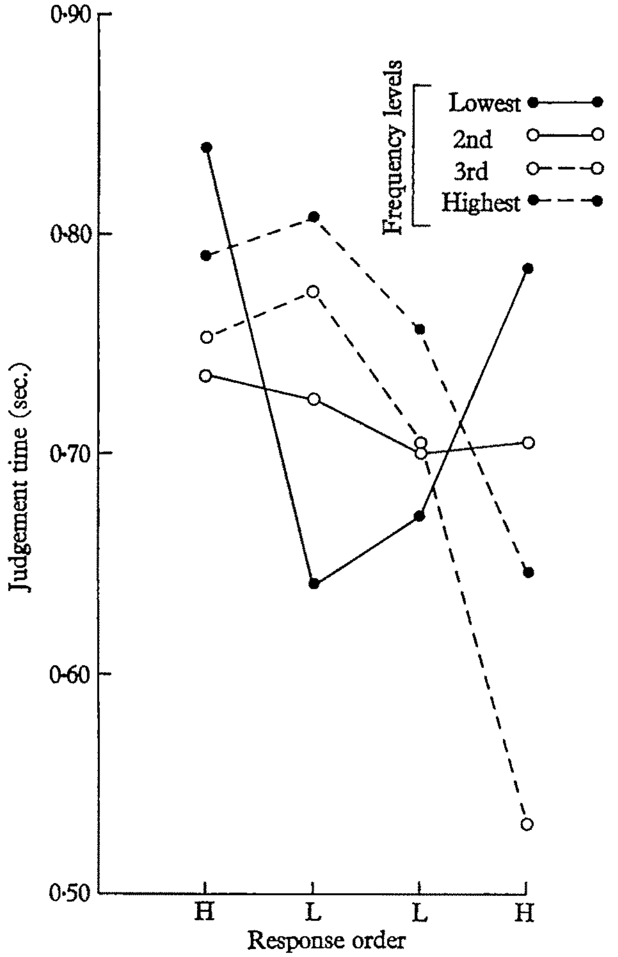


Fig. 5. Practice effects, 1. Mean judgement times at each frequency level for blocks of 'which higher' and 'which lower' responses in series order.

Variations in the curves caused by differences in the relative ease of the four discriminations should be discounted. Such variations would cause a movement of a whole curve representing the results for a given frequency level up or down the ordinate scale. Thus the absence of any intersection of the curves on the response order graphs is not evidence against the 'cross-over' effect. This effect is seen in the opposing direction of the movements of high and low frequency curves when progressing along the abscissa from H to L or from L to H response conditions.

Stimulus order effects

On each presentation the subject heard two stimuli, differing in pitch, and separated by a short interval. He was required, on different occasions, to choose the higher or the lower. On half the presentations the stimulus chosen was the first of the pair, and on half the second. To see whether the cross-over effect was influenced by this temporal ordering of the stimuli the results were analysed for the cases when the stimulus tone corresponding to a correct judgement was the first tone and then separately for when it was the second tone that matched the required judgement.

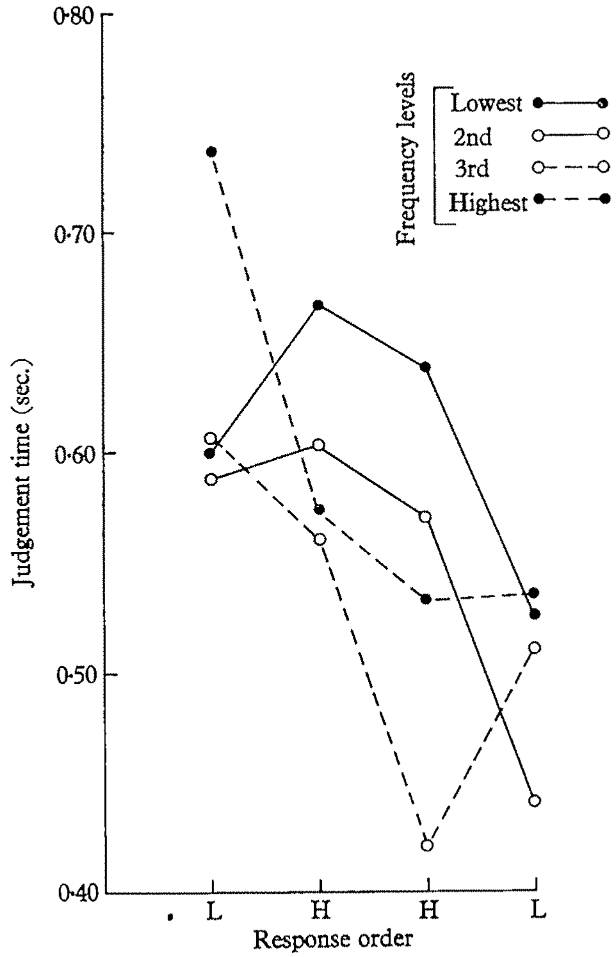


Fig. 6. Practice effects, 2. Mean judgement times at each frequency level for blocks of 'which lower' and 'which higher' responses in series order.

Fig. 7 shows the influence of stimulus order on the cross-over effect to be striking and consistent. At every frequency level and for both response conditions the cross-over is markedly exaggerated when it is the second stimulus that matches the correct judgement. When the first stimulus is chosen the cross-over is reduced, and for the highest frequency level is reversed. Thus in this type of sequential judgement, with the stimuli separated in time, four different factors appear to interact and jointly affect the speed of judgement: (1) the response instruction; (2) The relative magnitude of the stimuli; (3) the direction of the change of pitch moving from the first to the second stimulus; (4) whether the stimulus responded to is the first of the pair or the second. For example with a response set to choose the lower tone and on the presentation of a low initial stimulus, when this is followed by an even lower comparison stimulus an extremely short latency is achieved. The term response congruity seems apt for the description of this relationship between the stimuli forming the pair to be discriminated and the nature of the response required. However, an examination of the various combinations of response congruity and the other factors did not reveal any simple rules that will consistently cover all the findings and help to elucidate the judgemental processes involved.

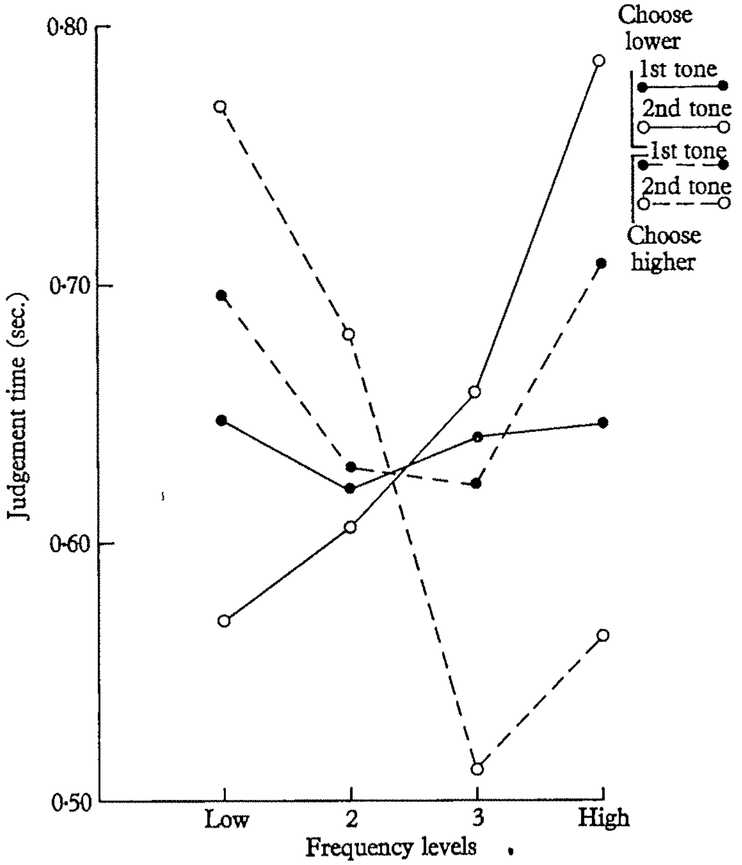


Fig. 7. Tone order effect. Mean judgement times at each frequency level for each response set when the tone responded to was first or second of comparison pair.

DISCUSSION

It has been demonstrated that the speed of pitch judgements is affected by an interaction between the response the subject is using and the frequency of the vibrations of the stimulus. The previous paper reported a similar 'cross-over' effect for brightness discrimination, and discusses the theoretical implications involved. Two explanations seem equally feasible. One might suppose that, for the slower stimulus response conditions, judgements are initially made in terms of the response category compatible with the general stimulus level, and are then translated into the mode of judgement demanded by the instructions. Thus low frequency tones will induce in the subject a tendency to judge in terms of lowness, necessitating a subsequent reversal of the terms of their discrimination on occasions where they have been told to choose which is the higher. The time taken by this reversal will therefore have to be added to the basic judgement for all the incompatible stimulus-response categories.

This explanation could be developed from a model of choice behaviour outlined elsewhere (Audley, 1960). Underlying the process of selecting from among a number of alternatives there is pictured a system generating covert responses representing these possible choices. The conflict between covert responses is terminated by the occurrence of an uninterrupted sequence of a certain number of covert responses representing a given alternative, which is then overtly chosen. An overt response is therefore likely to occur quickly the more rapidly these implicit response tendencies are generated. Now when a series of stimuli varying along a dimension is presented to a subject and the subject is instructed to adjudge these stimuli in categories of response relevant to this dimension, it is assumed that the magnitude of the stimulus will affect the rate at which the implicit responses are generated. In pitch dimension, the occurrence of a high tone will lead to a rapid rate of generation of implicit 'high' responses, and a slow rate of 'low' ones, the opposite state of affairs resulting from the occurrence of a low tone. However, there is an alternative to the assumption that the dominant response system is used and that the subject then translates the result of the judgement into the terms dictated by the experimenter. It could be that the subject chooses on the basis of the covert response system corresponding to the experimenter's instruction.

It follows that when two high tones are presented from which the subject is required to select the higher then the judgement time will be short, as the system will be generating implicit 'high' responses at a rapid rate. The selection of the lower of two high tones will tend to be slower, as the choice will be governed by a less rapid implicit response system.

If the nature of the cross-over effect for stimulus levels intermediate in a series could be discovered it was felt that this might help to elucidate the mechanisms involved. The results obtained with 4 pitch levels suggest that the cross-over is graded over the middle stimuli, a choice of 'which higher' being increasingly facilitated as the pitch level moves higher up the series. At first sight this might appear to support the notion of an underlying graded effect to argue against the response translation hypothesis.

A similar conclusion, however, could still be arrived at by a slight modification of this viewpoint. In introducing the idea it was assumed that, on receiving a pair of high tones, the subject was induced to choose the higher, and when required to respond in terms of 'which lower', a time lag resulted because the mode of judgement had to be reversed. Over a range of stimuli this would seem to suggest an all-or-none effect. But it would seem more likely that the operation of this mechanism would be probabilistic in nature. The higher the frequency of vibration, the more probable will it be that the initial response of 'which higher' will be invoked, though for only moderately high frequencies a number of initial 'lower' responses will also occur, this proportion increasing as the pitch series is descended. Thus for stimulus 2, under the response instruction of 'which lower' a smaller proportion of judgements will have to be translated than judgements of stimulus 3 under the same instruction, though a larger proportion than for stimulus 1. Again a graded cross-over effect would be predicted. However, if this is the situation, then the variability of response times for any given response should be greater for the two intermediate tones.* This would occur because in the middle of the pitch series the subject will use both response sets about equally often and hence will have an extreme mixture of fast and slow times. Using the range as a rough measure of variability, then going from the lowest to the highest pitch, the average ranges for observations within a given response instruction are 0·30, 0·43, 0·29 and 0·39 secs. This casts some doubt on the translation hypothesis.

Without a great deal of elaboration it is not possible to relate the results of the present study to earlier ones concerned with contextual effects in psychophysical situations. One major difficulty is that the work reported here used time measures in an almost error-free situation whereas the earlier experiments studied changes in the frequency with which different response categories are used. Nevertheless, it is worth noting that the cross-over effect resembles the phenomenon of 'central tendency'. Response errors leading to 'central tendency' result from too many judgements of 'lower' to a low pair of stimuli, and too many judgements of 'higher' to a high pair of stimuli. Using the method of constant stimuli, such central tendency errors have been reported for judgements of pitch (Wolfe, 1886; Wada, 1932; Koester & Schoenfeld, 1946), for judgements of brightness (Lewis, 1909), of loudness (Needham, 1935), and lifted weights (Woodrow, 1933).

REFERENCES

AUDLEY, R. J. (1960). A stochastic model for individual choice behaviour. *Psychol. Rev.* **67**, 1–15.

AUDLEY, R. J. & WALLIS, C. P. (1964). Response instructions and the speed of relative judgments. I. Some experiments on brightness discrimination. *Brit. J. Psychol.* **55**, 59–73.

KOESTER, T. (1945). Time-error and sensitivity in pitch and loudness discrimination as a function of time interval and stimulus level. *Arch. Psychol.* no. 297.

KOESTER, T. & SCHOENFELD, W. N. (1946). The effect of context upon the judgments of pitch differences. *J. Exp. Psychol.* **36**, 417–30.

LEUBA, J. H. (1892–93). New instruments for Weber's Law: indications of the law of sense memory. *Amer. J. Psychol.* **5**, 370–84.

LEWIS, T. A. (1909). On the transformation of memory content in the comparison of lights. *Psychol. Rev. Mon. Suppl.* **10**, whole no. 40, 55–62.

NEEDHAM, J. G. (1935). The effect of the time interval upon the time error at different intensive levels. *J. Exp. Psychol.* **18**, 530–43.

POSTMAN, L. (1946). The time error in auditory perception. *Amer. J. Psychol.* **59**, 193–219.

STEVENS, S. S. (1957). On the psychophysical law. *Psychol. Rev.* **64**, 153–81.

STEVENS, S. S. & VOLKMANN, J. (1940). The relation of pitch to frequency; a revised scale. *Amer. J. Psychol.* **53**, 329–53.

WADA, Y. (1932). Der Zeitfehler beim Sukzessivvergleich der Tonhöhen. *Jap. J. Psychol.* **7**, 505–38.

WHIPPLE, G. M. (1901). Analytic study of the memory image and the process of judgment in the discrimination of clangs and tones. *Amer. J. Psychol.* **12**, 409–27.

WOLFE, H. K. (1886). Untersuchungen über das Tongedächtnis. *Philos. Stud. (Wundt)*, **3**, 534–71.

WOODROW, H. (1933). Weight discrimination with a varying standard. *Amer. J. Psychol.* **45**, 391–416.

---

* This was drawn to our attention by T. Shallice. An experiment involving variations in the difficulty of the discriminations is being conducted to distinguish between these two alternative explanations.

(*Manuscript received* 6 *March* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 133–141
Printed in Great Britain

CONSTANT ERRORS IN WEIGHT JUDGEMENTS AS A FUNCTION OF THE SIZE OF THE DIFFERENTIAL THRESHOLD

BY HELEN E. ROSS
Psychological Laboratory, University of Cambridge

Constant errors in psychophysical procedures are known to increase in magnitude with the intensity of the standard, and with the temporal or spatial separation of the standard and comparison stimuli. It is often assumed that this is due to increasing 'adaptation' of the effective standard. It is argued here that the increase is due to a statistical artifact, since the measures of constant error and of differential threshold normally used are not independent, so that any factor which increases the difficulty of discrimination will increase the magnitude of the constant error.

It is shown that for differential thresholds ranging from 1 to 25 g. the absolute size of the constant error increases with the size of the threshold, but the direction of the error is variable and remains unexplained.

I. INTRODUCTION

In psychophysical experiments it is often found that a difference arises between the 'point of subjective equality' (that stimulus which is judged equal to the standard) and the physical value of the standard. This difference is called the 'constant error' (CE). CE's most frequently arise in comparison methods, when the standard is distinguishable from the comparison stimulus because it is always presented first (or second), or on the left (or right). With temporal separation the CE is usually called the 'time error' and with spatial separation the 'space error'.

The present investigation was undertaken as a result of a previous experiment by Ross & Gregory (1964) in which it was shown that the size-weight illusion influenced the differential threshold (DL): when the physical weight was kept constant apparently heavy weights gave rise to higher DL's than apparently light weights. In this experiment the Method of Constant Stimuli was used, with a standard of 200 g., and the comparison weights were lifted either simultaneously or before or after the standard. It was also observed that larger CE's were associated with the apparently heavier weights, and that there was a correlation between the DL and the absolute size of CE. (The modulus sign, $|\text{CE}|$, will be used to indicate that only the size and not the direction of CE is being considered.) This correlation is to be expected since the measures of CE and DL normally used are not statistically independent. However, some authors have considered the CE to be an independent variable due to some kind of adaptation, and if this were so the adaptation might affect the DL and thus give rise to a correlation between $|\text{CE}|$ and DL apart from the statistical artifact.

Köhler (1923) suggested that the 'time error' was due to a 'fading trace' left by the first stimulus with which the second stimulus was compared and overestimated. However, explanations based on 'fading' assume that the trace does not merely become disturbed with time, but represents a signal of lower intensity: they therefore fail to explain the fact that the direction of the error is sometimes to *underestimate* the second weight (Woodrow, 1933). Moreover, time errors also occur in judgements of quality (as opposed to quantity or intensity) where 'fading' could not apply. For example, Koester & Schoenfeld (1946) found time errors in pitch discrimination.

These authors unfortunately give no data on the effect of the time interval upon the DL, so we cannot be certain that they were increasing together. However, Piaget & Lambercier (1943) include a measure of the DL in their work on the 'space error' in the comparison of lines: in this case the DL increases with the spatial separation of the lines, and adequately accounts for the increase in the CE.

According to Weber's Law the DL increases with the intensity of the standard, and there is considerable evidence showing that the size of the CE also increases. It is usually found that comparison stimuli are overestimated in relation to intense standards and underestimated in relation to weak ones, but that the absolute size of the error increases with the intensity of the standard. Hollingworth (1910) proposed the theory of the 'central tendency of the standard' to explain both the size and the direction of the CE: he supposed that the value of the standard in some way gravitated towards the mean of the series under consideration, with the result that there was no CE with standards at the mean but large CE's in opposite directions with standards at the extremes of the series. It is not well established that there are large errors of underestimation at low intensities, though overestimation at high intensities has been shown in several modalities: for example, by Hollingworth (1910) for judgements of length, Woodrow (1933) for weight, and Needham (1935) for loudness. The errors are normally asymmetrical, with considerably more overestimation than underestimation, so Woodrow suggested that in addition to 'central tendency' there was a continual tendency for the value of the standard to sink with time. Helson (1947) proposed a similar but more complex 'adaptation level' theory, in which the value of the effective standard gravitated towards a weighted geometric mean of the series, with a negative constant added for the 'time error'. These theories are unsatisfactory because they assume the occurrence of a negative 'time error' of constant size without proposing any method of measuring it independently of 'central tendency', or allowing for the correlation between the DL and the magnitude of any measure of constant error.

It seems advisable to treat the direction and the magnitude of the CE as two separate questions. Most of the 'central tendency' data are presented as group means taking the direction of error into account; since individual subjects vary in the direction of their error, the underlying Weber relationship is often disguised, especially with weak standards when subjects may change from overestimation to underestimation. The only data inconsistent with Weber's Law were those of Needham (1935) on loudness discrimination, where individuals gave increasingly large CE's at low intensities. However, we cannot be certain that this result conflicts with Weber's Law, since no data on the DL were included, and an unstable measure of CE was used. This measure was $D\,\%$, which is found by taking the difference between the number of judgements 'heavier' and 'lighter', divided by the total number of judgements and expressed as a percentage. Since 'equal' judgements are allowed, the size of $D\,\%$ will vary inversely with the number of such judgements. in addition to increasing with the DL.

Since it was not clear from the existing data whether the CE was proportional to the standard at low intensities it was decided to perform an experiment on weight discrimination similar to the previous experiment of Ross & Gregory but using much lighter weights. Only one time order was used, as this was not intended as a study of all the different variables which might affect the direction and size of the error.

II. METHOD

Eight subjects were used. Seven were undergraduates and graduates of both sexes, aged between 19 and 29 years; one was a boy of 13.

Three sets of weights were used, each set consisting of one standard and nine comparison weights. The weights were made from tins containing lead shot and candle wax, and fitted with wire loops on the lids with which to lift them. All the tins were painted black. One set was made from Nescafé tins 3 in. high by 3 in. diameter, one from Kodak film tins 1¾ in. by 1¼ in., and one from Ferrania film tins 2¼ in. by ¾ in. The smallest tins (Ferrania) ranged from 6 to 14 g. in 1 g. intervals, with a standard of 10 g. equal to the middle weight. The medium-sized (Kodak) and large-sized (Nescafé) tins both ranged from 65 to 85 g. in 2·5 g. intervals, with a standard of 75 g. The only difference between these two sets of weights was size, with the effect that the smaller set felt heavier (size-weight illusion). The purpose of this was to attempt to obtain results at a lower intensity comparable to the previous results of Ross & Gregory. In that experiment the size-weight illusion was shown to influence the DL and CE, and in this experiment it was hoped to discover whether the same effect occurred at low intensities.

The subject and the experimenter were seated at a table with a screen between them. Each subject was tested on all three sets of weights, the order of sets being varied for different subjects, but the same procedure used for each set. The subject was instructed to use his preferred hand throughout the experiment, and to lift first the standard and then the comparison weight. The comparison weights were hidden behind the screen, and were presented to the subject in a predetermined random order, each weight appearing ten times. The subject lifted the weights up and down smartly by the loop once, and then said whether he thought the comparison weight was heavier or lighter than the standard. 'Equal' judgements were not allowed; if uncertain the subject had to guess.

III. RESULTS

The standard deviation (SD) and constant error (CE) for each subject and each set of weights were calculated according to the method of averaged z-scores, as described by Woodworth & Schlosberg (1955, p. 205). In this method the probability is calculated of judging each comparison weight heavier than the standard, and each probability score is transformed into a z-score (normal deviate score). It is assumed that the z-scores are linearly related to the weight of the stimulus, and a straight line is fitted through two points which are found by averaging the z-scores for the lighter half and the heavier half of the comparison weights separately. The SD is a measure of the slope of this line, and is found by calculating the number of grams on the x-axis corresponding to one standard deviation on the y-axis. The CE is given by the difference between the weight of the standard and the weight of the comparison stimulus judged heavier on 50 % of the trials. The response bias (RB), or the degree of overestimation or underestimation of the standard, was also calculated. This measure is given by the probability of judging the comparison weight as heavier when it is equal to the standard, subtracting 0·5. Thus CE and RB are functionally related variables, since CE is the intercept on the x-axis and RB the intercept on the y-axis of the same response curve at $p = 50$ and $x = S$ (standard) respectively. Fig. 1 shows typical response curves for subjects with little error and with positive and negative errors. The y-axis here shows probability scores in place of the corresponding z-scores, but so spaced as to give the same straight-line transformation for the response curves.

As a measure of correlation Kendall's τ was calculated for each condition separately between SD and the absolute size of CE and RB (|CE|, |RB|). The correlations are shown in Table 1. In every condition there was, as expected, a strong correlation between |CE| and |RB| ($P < 0{\cdot}001$); but the correlation between |CE| and SD was never significant, and that between SD and |RB| was even smaller and in some cases negative. However, when there is a sufficiently large variation in SD the correlation between SD and |CE| becomes significant. This is shown graphically in Fig. 2, where the results for individual subjects are shown from all the experiments in which the standard was lifted first. The variation in the ranges of weights used produces a large variation in SD, and the correlation is significant at the 1 % level on a graphical test of monotonic association (Quenouille, 1959, no. 28). The same results are shown

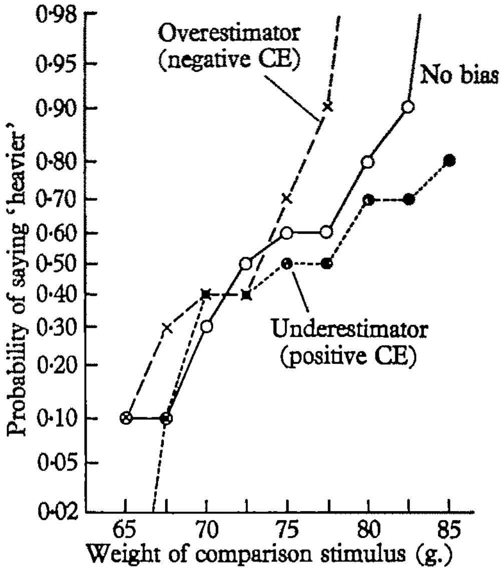


Fig. 1. Typical response curves for three subjects using the large tins and a standard of 75 g. Probability graph paper is used, so that if the response distributions are normal they should be fitted by straight lines.

Table 1. *The Kendall rank correlation coefficient between SD, |CE| and |RB| under different conditions*

(The correlation between |CE| and |RB| is always significant ($P < 0{\cdot}001$ two-tail test), but the correlation between SD and |CE| or |RB| is never significant.)

Condition	Weights	Weight of standard (g.)	No. of subjects	SD and |CE|	SD and |RB|	|CE| and |RB|
Standard first	Small	200	16	0·167	0·017	0·834
Standard first	Large	200	16	0·050	−0·017	0·860
Comparison first	Small	200	16	0·227	−0·113	0·700
Comparison first	Large	200	16	0·133	−0·017	0·683
Simultaneous	Small	200	20	−0·100	−0·116	0·759
Simultaneous	Large	200	20	0·290	0·147	0·854
Standard first	Small	75	8	0·097	0·032	0·387
Standard first	Large	75	8	0·097	0·0	0·353
Standard first	Very Small	10	8	0·161	0·129	0·404

in Fig. 3, except that |RB| is used instead of |CE|. A similar trend is present as in Fig. 2, but the correlation just fails to be significant at the 5 % level on the same test.

Since |RB| is less affected by SD than is |CE| it is used in Fig. 4 to show the distributions of errors obtained under different conditions. The first three pairs of histograms are derived from the data of the previous experiment, and the last three histograms from this experiment. The sign of RB is the opposite of that of CE: thus overestimators have a positive RB but negative CE, and vice versa. The medium SD

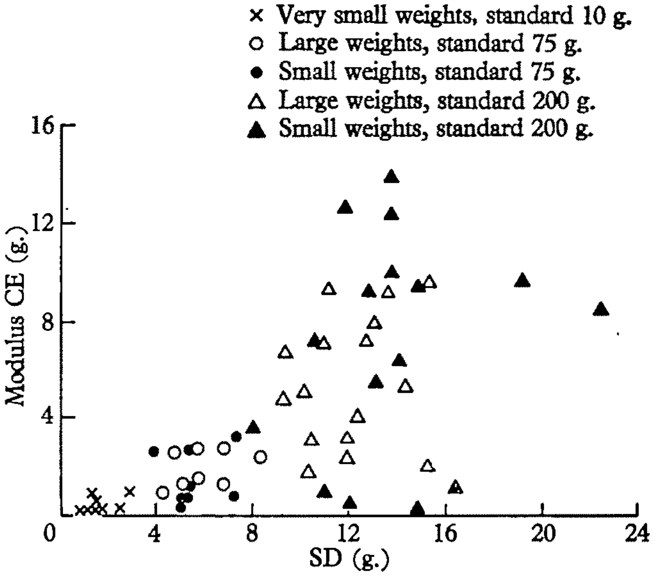


Fig. 2. Scatter diagram showing a correlation between SD and |CE|. The data are taken from all the conditions in which the standard was lifted first.

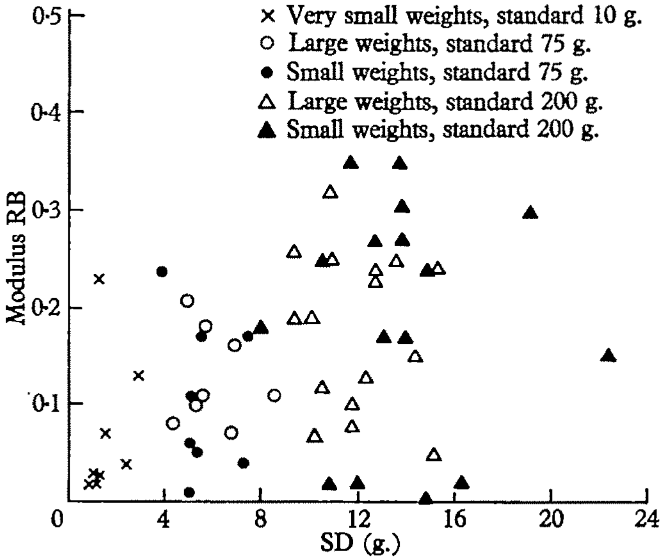


Fig. 3. The same data as in Fig. 2 except that |RB| is used instead of |CE|. The correlation is no longer significant.

for each group is shown, and it can be seen that in the previous experiment there are both larger SD's and larger RB's than in the present experiment. There are also changes in the direction of RB under different conditions. In general there is overestimation with the standard lifted first, underestimation with the comparison stimulus lifted first, and slight overestimation with simultaneous lifting. There is also a tendency for the size-weight illusion to affect the direction of the error: when the small (Kodak) tins are compared with the large (Nescafé) tins it can be seen that more subjects show errors in the predominant direction, producing a shift in the means of the distributions. This difference only reached a significant level in the previous experiment in the case where the standard was lifted first ($P = 0{\cdot}01$, Wilcoxon test). In the present experiment no significant difference in either SD or CE was found between these two sizes of tins, perhaps because too few subjects were used.

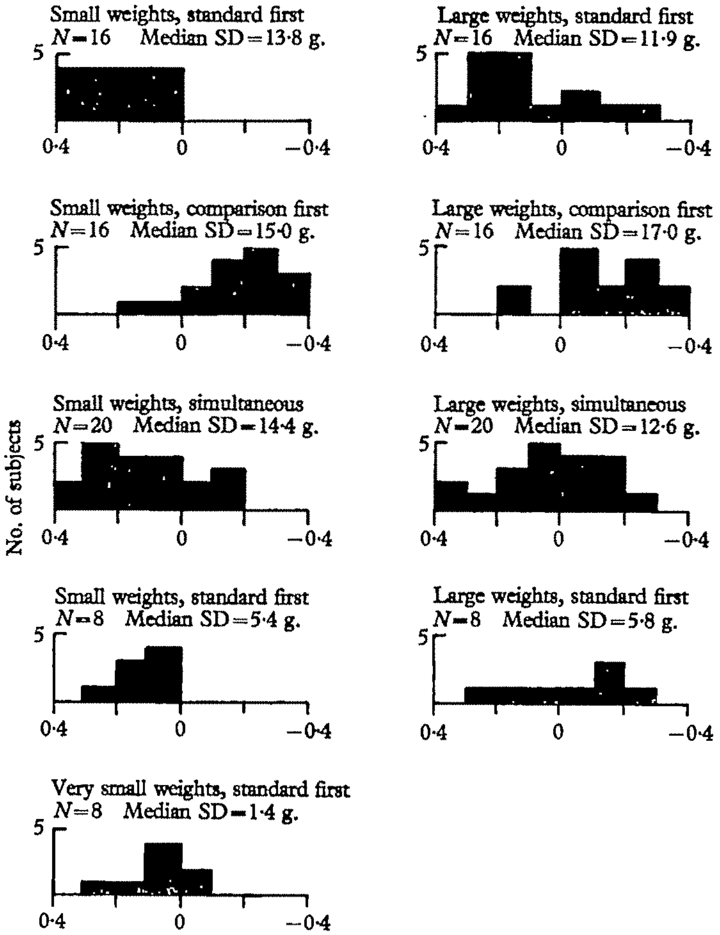


Fig. 4. Histograms showing distributions of positive and negative response biases obtained under different conditions. The first six are from the previous experiment, the last three from this.

IV. DISCUSSION

The method of averaged z-scores may result in slightly inaccurate estimates of SD and CE when the CE is large, since the method is based on the assumption that the responses form a normal cumulative probability curve, with the point of subjective equality (PSE) in the centre. For large CE's the PSE is not central, and the distributions do not appear to be normal. The inaccuracy of this method is small compared with the observed correlation between $|\text{CE}|$ and SD, but it suggests that the assumption of normality should not be made in calculating the results. Thurstone (1928) suggested that in view of Weber's Law the response distributions should be plotted on a logarithmic scale. This has little effect over the small range of weights used in these experiments. Moreover, this would only account for positive CE's, whereas CE's can occur in either direction producing response curves which are symmetrically skewed: overestimators produce curves which rise more steeply over the heavier than the lighter weights, and the converse is true for underestimators (see Fig. 1).

Regardless of the question of the most suitable method of calculating the constant error and differential threshold, the pattern of results is fairly clear: $|\text{CE}|$ is a function of both $|\text{RB}|$ and SD, and in any set of data it correlates most highly with the factor showing the larger fluctuations. In these experiments SD was varied by using different subjects, weights, sizes of tin and methods of presentation. In any one condition the variation in SD among different subjects was not great enough to give rise to a significant correlation, but when the intensity of the weights was also varied the correlation became significant. $|\text{RB}|$ is not significantly correlated with SD, though with large variations in SD it shows a small positive correlation due to the non-independence of the measures. In future investigations of the factors affecting the size of the constant error it would be advisable to use $|\text{RB}|$ as a measure, and to reduce variation in SD as far as possible.

Changes in the direction of CE are much harder to explain than changes in size. The direction is known to be influenced by the temporal or spatial order of stimuli, the size of the temporal or spatial interval, the intensity of the standard, and by practice (Needham, 1934). The CE is notoriously unstable, and can be affected by the form of question and response (Brown, 1910). In the present experiment there was a change to more positive CE's with the larger (and apparently lighter) Nescafé tins as opposed to the Kodak tins, but this trend was not continued with the small and very light Ferrania tins. This does not necessarily contradict Woodrow's results, as he obtained positive errors for light standards only when the weight of the standard was varied throughout the experiment instead of being kept constant for each series. Moreover, he used the same size of tin for all the weights, which would enhance the lightness of the light weights (size-weight illusion). Perhaps some kind of 'contrast' is necessary (whether due to the size-weight illusion or to other weights in the series) before positive errors are obtained with the standard presented first.

Explanations in terms of peripheral adaptation (e.g. Helson, 1948) hardly seem appropriate, since CE's occur in such a variety of modalities and conditions of presentation that it is difficult to postulate suitable adaptation mechanisms for all of these (e.g. the simultaneous presentation of lines of different lengths). Whether adaptation is thought to be peripheral or central it is implied that the CE is an independent variable due to some perceptual change. If this were the case there should be a negative correlation between SD and |RB|, since the PSE is regarded as fixed, and a flatter slope of the response curve must make a smaller intercept on the y-axis, and thus a smaller |RB|. The results do not support this conclusion, but suggest instead that |CE| is a function of |RB|, with SD acting as a scaling factor: we should regard |RB| as fixed, so that a flatter slope of the response curve will make a larger intercept on the x-axis, and thus a larger |CE|.

It has been pointed out by Irwin (1958) and Luce (1959, pp. 30–4) that subjects may show a preference for one response which is unrelated to the ability to discriminate or solve problems. A response bias of this sort might be considered similar to the gambling strategy described by Bruner, Goodnow & Austin (1956), when subjects were asked to sort different pictures of aeroplanes into two categories. With no knowledge of results they tended to pick on an irrelevant cue and use it persistently with little regard for the outcome; while with some knowledge of results they made an effort to categorize correctly, and tended to guess in the proportion in which they thought the two categories were occurring. Psychophysical tasks are similar in that they demand categorization with incomplete knowledge of results. This suggests that CE's might represent a form of gambling in favour of one response when uncertain: the gamble could be completely haphazard or it could be in favour of the category believed to be more frequent or more desirable. CE's do not seem to be haphazard since there is some regularity in their direction; nor do most subjects believe that the categories are unequally distributed—indeed, they are frequently disturbed by the apparent inequality introduced by their CE's; nor is there any obvious reason why one category should be regarded as more desirable. Possibly one response is favoured by verbal conditioning: we say 'heavier or lighter' 'louder or softer' rather than the other way round, and this may produce a preference for the first response (R. J. Audley, personal communication, 1962). Unfortunately this provides no explanation for errors in the reverse direction.

Explanations in terms of response bias assume that the subject is guessing rather than discriminating. In the statistical decision theory of Tanner & Swets (1954) it is assumed that there is almost no limit to the subject's ability to discriminate between stimuli of different intensities, but that the perceived intensity of a given stimulus is variable. The subject's task is to decide whether a given perceived intensity is more likely to have arisen from one or another physical stimulus. In making this decision he is influenced not only by the relative probabilities involved but by the different rewards and penalties for success or failure in either decision. In view of these he adopts a 'criterion' or 'cut-off point' for sorting the perceived stimuli into two categories. In the case of differential thresholds the standard is given as the criterion, and there seems to be no reason why CE's should arise unless the rewards and penalties in the two categories are sufficiently unequal to make the subject adopt some other criterion. As already stated, it is difficult to see why this should be so. However, experiments in which the probabilities and values were deliberately altered might reveal how far these factors influence CE's.

Decision theory is probably too precise in supposing that subjects are always discriminating rather than guessing. Senders & Sowards (1952) found that subjects tended to respond in short repetitive runs rather than random sequences. This suggests that they are not always maintaining a consistent decision rule, since each response is partly determined by the previous responses. Luce (1963) quotes evidence in favour of a discrimination threshold and suggests a threshold model incorporating response biases. In practice it is very difficult to distinguish between guessing and discriminating at threshold level, though there is a theoretical difference. Analysis of response sequences may reveal more about the nature of response biases than can be found from cumulative response distributions.

This work was carried out while supported by the U.S.A.F. European Office of Aerospace Research, Grant No. A.F. EOAR 63-93. I am grateful to Mr R. L. Gregory, Miss Violet Cane and Mr D. R. J. Laming for much valuable discussion and advice, and to Prof. O. L. Zangwill for providing facilities.

REFERENCES

BROWN, W. (1910). The judgement of difference. *Univ. Calif. Publ. Psychol.* **1**, no. 1.

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*, chap. 7. New York: Wiley.

HELSON, H. (1947). Adaptation-level as a frame of reference for prediction of psychophysical data. *Amer. J. Psychol.* **60**, 1–29.

HELSON, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.* **55**, 297–313.

HOLLINGWORTH, H. L. (1910). The central tendency of judgement. *J. Phil. Psychol. Sci. Meth.* **7**, 461–9.

IRWIN, F. W. (1958). An analysis of the concepts of discrimination and preference. *Amer. J. Psychol.* **71**, 152–65.

KOESTER, T. & SCHOENFELD, W. N. (1946). The effect of context upon judgement of pitch differences. *J. Exp. Psychol.* **36**, 417–30.

KÖHLER, W. (1923). Zur Theorie des Sukzessivvergleichs und der Zeitfehler. *Psychol. Forsch.* **4**, 115–75.

LUCE, R. D. (1959). *Individual Choice Behavior*. New York: Wiley.

LUCE, R. D. (1963). A threshold theory for simple detection experiments. *Psychol. Rev.* **70**, 61–79.

NEEDHAM, J. G. (1934). The time-error as a function of continued experimentation. *Amer. J. Psychol.* **46**, 558–67.

NEEDHAM, J. G. (1935). The effect of the time interval upon the time error at different intensive levels. *J. Exp. Psychol.* **18**, 530–43.

PIAGET, J. & LAMBERCIER, M. (1943). Recherches sur le developpement des perceptions. 2. *Arch. Psychol., Genève*, **29**, 123–253.

QUENOUILLE, M. H. (1959). *Rapid Statistical Calculations*. London: Griffin.

ROSS, H. E. & GREGORY, R. L. (1964). Is the Weber fraction a function of physical or perceived input? *Quart. J. Exp. Psychol.* (in press).

SENDERS, V. & SOWARDS, A. (1952). Analysis of response sequences in the setting of a psychophysical experiment. *Amer. J. Psychol.* **65**, 358–74.

TANNER, W. P. & SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychol. Rev.* **61**, 401–9.

THURSTONE, L. L. (1928). The phi gamma hypothesis. *J. Exp. Psychol.* **11**, 293–305.

WOODROW, H. (1933). Weight discrimination with a varying standard. *Amer. J. Psychol.* **45**, 391–416.

WOODWORTH, R. S. & SCHLOSBERG, H. (1955). *Experimental Psychology*. London: Methuen.

(*Manuscript received* 3 *January* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 143–154
Printed in Great Britain

CUES THAT ELICIT ANALYTIC-DEDUCTIVE METHODS IN CONCEPT ATTAINMENT*

BY WAYNE A. WICKELGREN
Massachusetts Institute of Technology

In inference concept attainment some subjects draw conclusions about concepts (global-hypothesis method) and some subjects draw conclusions about the relevance or irrelevance of dimensions (analytic-deductive method). Experiments were performed to determine whether the presence or absence of certain cues accompanying inference concept attainment problems would affect the incidence of analytic-deductive methods in these problems. Two such cues were identified. One effective cue was a series of three 'informational-analysis' questions which asked the subject what information he had, whether he had enough information to guess the concept, and, if not, what information he needed. The other effective cue was the introduction of multiple-choice answers using only analytic-deductive terms like 'relevant' and 'irrelevant', but permitting both global-hypothesis conclusions and analytic-deductive conclusions to be drawn using these terms. Both of the cues that elicited analytic-deductive conclusions to inference problems produced positive transfer to a selection concept attainment problem in which totally different responses were required.

I. INTRODUCTION

In an *inference* concept attainment problem subjects are given some examples and perhaps some non-examples of a concept. From this information they may be asked to determine the concept or to draw whatever conclusions they can about the concept. In a *selection* concept attainment problem, subjects are usually given one example of a concept. They are then permitted to select other instances and ask whether these instances are examples or non-examples. In this way they are to determine the concept as soon as possible.

The present experiment is concerned only with concepts that are *conjunctive* in form. The term 'conjunctive concept' will be used to refer to both *simple* and *and* concepts. Examples of a *simple* concept must meet one requirement (for example, red colour). Examples of an *and* concept must meet two or more requirements (for example, red colour and square shape). Subjects are instructed to assume that all concepts are conjunctive in form.

The instances in the present experiment are always six-place numbers. The dimensions of an instance are the six places of the number, and there are ten possible values for each dimension. Subjects are told what the dimensions and values are.

When the dimensions of the instances and the form of the concept are known, there are several possible methods for solving concept attainment problems. The methods can be categorized in two principal ways. First, does the method attack the problem as a whole (global method) or break up the problem into subproblems (analytic method)? A global approach to a concept attainment problem means trying to determine the concept. An analytic approach to a concept attainment problem means trying to determine the relevance or irrelevance of dimensions. Secondly, does the method involve formulating hypotheses that are consistent with the information, but not logically implied by the information (hypothesis method) or does the method involve only logical deductions from the information (deductive method)? This analytic scheme yields four types of problem solving methods for concept attainment: global-hypothesis, global-deductive, analytic-hypothesis, and analytic-deductive methods.

---

* Based on a thesis submitted in partial fulfilment of the requirements for the Ph.D. degree from the University of California, Berkeley. The research was facilitated by a grant from the Carnegie Corporation of New York to Richard S. Crutchfield at the Institute of Personality Assessment and Research, University of California, Berkeley.

These four types of methods are slightly different from the concept attainment strategies suggested by Bruner, Goodnow & Austin (1956), but the relationship is close. These authors suggested two types of strategies for inference concept attainment problems: wholist and partist strategies. As Bruner *et al.* defined these strategies, both of them would be global-hypothesis methods since they both involve formulating hypotheses about the entire concept. These two strategies for inference concept attainment are clearly inadequate because they do not include any strategy that determines relevant dimensions from the information. Even in an experiment where subjects were forced to formulate hypotheses about the concept on each trial, Bruner *et al.* reported only a little better than 50 % adherence to the rules of the strategy a subject was classified as using.

Bruner *et al.* suggested four strategies for selection concept attainment problems: *simultaneous scanning*, *successive scanning*, *conservative focusing*, and *focus gambling*. *Simultaneous scanning* is a global-deductive method: (1) because information is used to eliminate concepts, not to determine the relevance of dimensions, and (2) because only logical deductions are made from the information. *Successive scanning* is a global-hypothesis method because information is used to generate and test hypotheses about the concept that are consistent with the information, but not implied by it. *Conservative focusing* and *focus gambling* are both analytic-deductive methods because information is used to make deductions about the relevance or irrelevance of dimensions. If a subject formulates hypotheses about the relevance or irrelevance of particular dimensions (these hypotheses being consistent with the information, but not implied by it), then the subject is using an analytic-hypothesis method.

Bruner *et al.* report no use of global-deductive methods or analytic-hypothesis methods in their experiments on concept attainment. Subjects used only global-hypothesis methods or analytic-deductive methods. Very little is known about the variables that determine the relative frequency of global-hypothesis methods as against analytic-deductive methods in concept attainment problems. For selection problems Bruner *et al.* report that presenting all possible instances in a *regular* array (adjacent instances differing in one or a few dimensions only) produces more use of analytic-deductive methods than does presenting the instances in a *random* array. Bruner *et al.* report also that *abstract* materials (geometric figures with different colour, shape, number of figures, etc.) produce more use of analytic-deductive methods than do *thematic* materials (human figures with different sex, dress, facial expressions, etc.). Both of these findings are open to alternative interpretations (Wickelgren, 1962).

Many further distinctions can be made between different global-hypothesis or analytic-deductive methods for concept attainment, but these distinctions are not relevant to the following experiment. The primary purpose of the following experiment is to determine some of the variables that influence the relative frequency of analytic-deductive methods as against global-hypothesis methods in concept attainment. Two variables were selected for investigation, both involving transfer from different conditions of prior training.

One type of training, called informational-analysis training, was designed to provide the cue that the information in a concept attainment problem should be divided into two categories. The two categories were: (1) the information about the concept that was possessed by the subject after reading the problem, and (2) the information about the concept that was still needed by the subject to solve for the concept after reading the problem.

The other type of training, called multiple-choice training, involved forcing the subject to select one or more multiple-choice answers to a concept attainment problem. The multiple-choice answers all used the analytic-deductive terms 'relevant' and 'irrelevant', but it was quite possible to draw an analytic-hypothesis conclusion about the concept by specifying the relevance or irrelevance of every dimension.

The effect of these two types of training on the frequency of analytic-deductive as against global-hypothesis conclusions was determined by comparison with various control groups receiving different types of training. The concept attainment problems used in this experiment were 'binary inference problems' in which two examples or one example and one non-example of a concept were presented and subjects were asked to draw conclusions from the information, using a restricted language that permitted both global-hypothesis conclusions about the concept and analytic-deductive conclusions about the relevance or irrelevance of dimensions. The information presented in the binary inference problems was always *insufficient* to determine one and only one concept.

Transfer from the informational-analysis training to the multiple-choice training was investigated, and also transfer from inference problems to selection problems.

II. METHOD

(*a*) *Procedure*

The instances employed in the experiment were six-place decimal numbers. The dimensions were the six places of the number, and there were ten possible values for each dimension. Subjects received complete instruction on this matter.

The test problems were binary inference problems that differed in one or two places. Thus, there were four different types of binary inference problems: two examples that differed in one place ($2EC_1$), two examples that differed in two places ($2EC_2$), one example and one non-example that differed in one place ($EC\text{–}NEC_1$), and one example and one non-example that differed in two places ($EC\text{–}NEC_2$). In the test problems, a subject was merely directed to draw conclusions from the information. He was not told what sort of conclusions to draw, being left to formulate the proper questions to ask of this information and the proper answers to the questions he asked.

To ensure unambiguous verbal reports, subjects were required to construct all their answers from the terms of a *limited response language*. The language was developed by reference to previous research on concept attainment and by pretesting which used intuition, thinking aloud, and extensive probing to determine the sort of concepts of which subjects make frequent use. The response language was explicitly taught in a 'teaching machine program' by requiring a subject to solve problems in the correct use of each of the terms. This provided a rigorous behavioural criterion for the assertion that subjects knew the meaning of each term in the language. Inasmuch as the number of possible statements that can be constructed from this relatively limited, well-defined vocabulary is extremely large, we gain many of the advantages of constructed response while avoiding ambiguity in interpretation of statements. No assertion is made that the nature of the vocabulary and the nature of the vocabulary training have no effect on the conclusions drawn in binary inference problems. However, these effects are constant in this experiment and open to investigation in other experiments by manipulating the vocabulary or the vocabulary training.

This vocabulary was as follows:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9
P1, P2, P3, P4, P5, P6 = Place 1, ..., Place 6
And, And/Or, Or, is, are, in, if, then, whether
C = Concept   NC = Not concept
EC = Example of concept   NEC = Not example of concept
PC = Part of concept   NPC = Not part of concept
Rel = Relevant to concept   Irr = Irrelevant to concept

A binary inference problem looked like this:

The following are an example and a non-example of a Simple or And concept: 7 4 0 5 9 1 is EC. 7 4 0 5 3 1 is NEC. Your conclusions are?

An informational-analysis problem looked like this:

7 4 0 5 9 1 is EC. 7 4 0 5 3 1 is NEC. (For a Simple or And C.) What information do you have? Do you have enough information to guess C? (If you said yes) What is C? (If you said no) What information do you need?

Feedback was provided for only the second question; subjects were told they did not have enough information to guess the concept. This had the effect of forcing them to analyse information into two categories, without ever suggesting the appropriate way to do it.

A multiple-choice problem looked like this:

7 4 0 5 9 1 is EC. 7 4 0 5 3 1 is NEC. (For a Simple or And C.) Choose *one or more* of the following statements as correct logical inferences to be made on the above information: P5 is Rel. P1, P2, P3, P4, and P6 are Rel. P5 is Irr. P1, P2, P3, P4, and P6 are Irr.

All subjects received extensive language training in the meaning of C, EC, NEC, Pi, j in Pi, And, Or, And/Or, PC, NPC, Rel, and Irr. In addition, all subjects were given two ordinary inference concept attainment problems *without answers*. The ordinary inference problems presented examples and non-examples in sufficient number to determine the concept uniquely. Then all subjects were given the first test of the four binary inference problems. All 'tests' of the binary inference problems presented the four problems, once each. Following the first test of the binary inference problems, subjects assigned to six different groups received different problems. Three of the groups then received eight informational-analysis problems. The other three groups served as various control groups for the effects of the informational-analysis problems.

The *informational-analysis spontaneous-thinking control group* received 36 filler problems on the meaning of C, EC and NEC. These language instruction problems had previously been presented and were well learned. An already well-learned task was selected as most likely to give subjects opportunity for whatever spontaneous thinking about the test problems might occur. Any test-retest improvement is also measured by this control group.

The *informational-analysis problem-familiarization control group* received the eight binary inference problems given in the informational-analysis problems, in test form without the additional informational-analysis questions. This group was designed to determine what learning of analytic-deductive conclusions occurs as a result of doing a series of binary inference problems.

The *informational-analysis concept-guessing-denial control group* received the eight binary inference problems accompanied only by the question, 'Do you have enough information to guess the concept?' Feedback was provided for this question as in the informational-analysis problems. This group was designed to determine whether denying the global-hypothesis method of concept guessing was sufficient to reduce the frequency of global-hypothesis conclusions and increase the frequency of analytic-deductive conclusions.

Following the informational-analysis problems or the informational-analysis control problems, all three informational-analysis control groups and one of the three informational-analysis experimental groups were given 12 multiple-choice binary inference problems, the alternative answers being restricted to conclusions about the relevance and irrelevance of dimensions. A subject had to choose an answer in the analytic-deductive language, but he might draw a conclusion not justified by the information in the same manner that a guess of the concept is not justified by the information in a binary inference problem. That is, a subject might draw a conclusion about the relevance of all places just as one does implicitly in guessing a concept. The informational-analysis experimental group that received the multiple-choice problems is called the *experimental group*.

One of the remaining two informational-analysis experimental groups became the *multiple-choice spontaneous-thinking control group* receiving 12 filler problems on the meaning of C, EC and NEC, instead of the multiple-choice binary inference problems.

The other informational-analysis experimental group became the *multiple-choice problem-familiarization control group* receiving 12 binary inference problems in test form without multiple-choice answers.

Following multiple-choice binary inference problems or multiple-choice control group problems, all six groups were given a test of the binary inference problems, then eight multiple-choice binary inference problems with feedback as to the correct answers, then a final test of the binary inference problems and finally a selection concept attainment task to test transfer from inference to selection methods.

(*b*) *Subjects*

Eighteen subjects were randomly assigned to each control group, except for two groups which had 19. Thirty-one subjects were assigned to the experimental group.

The greater number of subjects assigned to the experimental group is statistically efficient since the experimental group is used in almost every comparison. Males and females were separately assigned to each group, and each group had almost exactly half males and half females. Subjects were University of California at Berkeley undergraduates enrolled in undergraduate psychology courses; they participated in the experiment as part of their course requirements.

Subjects were self paced, but the average time to completion (about 2 hr.) did not differ significantly among the six groups. Although each 2-hr. session included subjects in all six groups, almost no added instruction from the experimenter was necessary, and it was possible to administer the experiment to as many as 30 subjects at one time.

III. RESULTS

(a) *Types of conclusions to binary inference problems*

The correct analytic-deductive conclusion to a $2EC_1$ problem is that the place in which the digit was changed is irrelevant. The correct analytic-deductive conclusion to a $2EC_2$ problem is that both places in which digits were changed are irrelevant. The correct analytic-deductive conclusion to an $EC\text{–}NEC_1$ problem is that the place in which the digit was changed is relevant. The correct analytic-deductive conclusion to an $EC\text{–}NEC_2$ problem is that at least one (and possibly both) of the places in which digits were changed is relevant. If a subject drew a correct analytic-deductive conclusion to a binary inference problem, he was scored as using the analytic-deductive method.

A subject was scored as using a global-hypothesis method on a particular binary inference problem if he formulated a concept as his conclusion to the problem. These concepts were always consistent with the information given in the problem, but never implied by the information.

A few subjects in a few problems used the global-deductive method and defined the entire set of concepts consistent with the information. These occurrences were included in the totals for the correct analytic-deductive conclusions to avoid undue subdivision of the data.

Analogously, some subjects on some problems used the analytic-hypothesis method and drew conclusions about the relevance or irrelevance of every dimension, and thus, in essence, formed an hypothesis about the concept. These occurrences were included in the totals for the global-hypothesis method to avoid undue subdivision of the data.

Therefore, the two classes of methods for answering binary inference problems were *hypothesis methods* as against *deductive methods*. The terms 'global-hypothesis method' and 'analytic-deductive method' are retained in the discussion of the results to indicate that the overwhelming majority of all hypothesis conclusions were global, and of all deductive conclusions were analytic, in agreement with previous findings by Bruner *et al.* (1956).

There were two other categories of answers to binary inference problems: (1) incorrect analytic-deductive conclusions that were not specifications of an entire concept and (2) no conclusion in response to the problem. In general, the results were tabulated as *correct deductive conclusions* as against *all other conclusions* (including the above two categories of answers). In one the dichotomy was ***hypothesis conclusions*** as against *all other conclusions*.

(*b*) *What is learned from informational-analysis problems?*

Table 1 presents the number of subjects in the informational-analysis experimental and control groups who gave *correct deductive answers* to each of the four binary inference problems on the test *following* the informational-analysis experimental or control problems, of those subjects who answered the corresponding problem *incorrectly* on the test *before* these problems. Neither experimental nor control problems had any appreciable effect on subjects who answered the binary inference problems *correctly* on the first test.

Table 1. *Effect of informational-analysis problems on the frequency of correct deductive conclusions*

(Number of subjects giving D conclusions on T 2 of those subjects giving ND conclusions on T 1.)

	2 EC$_1$			2 EC$_2$			EC–NEC$_1$			EC–NEC$_2$		
Group	D	ND	*P*	D	ND	*P*	D	ND	*P*	D	ND	*P*
Informational-analysis experimental	21	30	—	23	29	—	17	28	—	13	47	—
Informational-analysis spontaneous-thinking control	2	11	—	3	9	—	1	9	—	1	15	—
Informational-analysis problem-familiarization control	1	13	*	1	12	*	1	13	*	1	16	—
Informational-analysis concept-guessing-denial control	1	14	*	1	14	**	2	16	*	1	16	—
Total informational-analysis control	4	38	***	5	35	**	4	38	**	3	47	*

D, correct deductive conclusion; ND, not correct deductive conclusion. T 1, first test of the binary inference problems; T 2, second test. P, * $< 0{\cdot}05$, ** $< 0{\cdot}01$, *** $< 0{\cdot}001$, one-tailed probability of difference from experimental group (Fisher's Exact Test).

The pattern of results in Table 1 is clear. The control groups are identical in performance and show almost no evidence of learning. In the experimental group, with the informational-analysis questions, roughly 40 % of the previously incorrect subjects learned to give correct deductive conclusions to the first three types of binary inference problems, and 25 % of the subjects learned to solve EC–NEC$_2$.

Table 2 pertains only to those subjects who gave hypothesis conclusions on the first test of the binary inference problems. It shows the number of these subjects who did and did not give hypothesis conclusions on the second test of these problems, which immediately followed the informational-analysis problems.

Of the subjects who were giving hypotheses in the 2 EC problems and the EC–NEC$_2$ problem, over 80 % abandoned hypothesis conclusions following the informational-analysis problems. For the EC–NEC$_1$ problem, the figure is 61 %. The control groups are not significantly different in abandoning hypothesis conclusions, but are significantly below the experimental group on every binary inference problem. Over the four binary inference problems an average of 20 % of the subjects in the three control groups abandoned hypothesis conclusions.

The informational-analysis problems increase correct analytic-deductive conclusions at the expense of incorrect global-hypothesis conclusions. The principal variable responsible for this effect is the cue of analysing information into two categories.

Spontaneous thinking, familiarization with binary inference problems, and denial of concept guessing are not sufficient to produce much learning of the analytic-deductive method or abandonment of the global-hypothesis method.

Table 2. *Effect of informational-analysis problems on the frequency of hypothesis conclusions*

(Number of subjects giving H conclusions on T 2 of those subjects giving H conclusions on T 1.

	2 EC$_1$			2 EC$_2$			EC–NEC$_1$			EC–NEC$_2$		
Group	NGH	GH	P	NGH	GH	P	NGH	GH	P	NGH	GH	P
Informational-analysis experimental	33	7	—	32	8	—	22	14	—	36	8	—
Informational-analysis spontaneous-thinking control	1	7	***	1	7	***	0	8	***	0	8	***
Informational-analysis problem-familiarization control	3	5	*	2	5	*	0	5	*	0	6	***
Informational-analysis concept-guessing-denial control	5	7	**	4	8	**	3	10	*	3	10	***
Total informational-analysis control	9	19	***	7	20	***	3	23	***	3	24	***

H, hypothesis conclusion; NH, not an hypothesis conclusion. T 1, first test of the binary inference problems; T 2, second test. P, * < 0·05, ** < 0·01, *** < 0·001, one-tailed probability of difference from experimental group (Fisher's Exact Test).

(c) *What is learned from multiple-choice problems?*

Table 3 reports the differences in learning of the deductive method among the experimental and multiple-choice control groups from the second to the third test of the binary inference problems. The second test occurred just prior and the third test just subsequently to the multiple-choice binary inference problems. Recall that all these subjects had received the informational-analysis problems just prior to the second test, and the data reported here are for only those subjects who were still *not* drawing correct deductive conclusions on the second test. The three types of problems (multiple-choice, constructed-response, or filler) presented after the second test of the binary inference problems had no differential effects on subjects who drew *correct* deductive conclusions on the second test.

The results in Table 3 are clear. The multiple-choice control groups are identical in

Table 3. *Effect of multiple-choice problems on the frequency of correct deductive conclusions*

(Number of subjects giving D conclusions on T 3 of these subjects giving ND conclusions on T 2)

	2 EC$_1$			2 EC$_2$			EC–NEC$_1$			EC–NEC$_2$		
Group	D	ND	P	D	ND	P	D	ND	P	D	ND	P
Experimental	7	4	—	7	4	—	7	3	—	9	9	—
Multiple-choice spontaneous-thinking control	0	9	**	0	9	**	0	10	**	0	17	***
Multiple-choice problem-familiarization control	2	7	**	2	7	—	0	8	**	1	12	*
Total multiple-choice control	2	16	**	2	16	**	0	18	***	1	29	**

D, correct deductive conclusion; ND, not correct deductive conclusion. T 2, second test of the binary inference problems; T 3, third test. P, * < 0·05, ** < 0·01, *** < 0·001, one-tailed probability of difference from the experimental group (Fisher's Exact Test).

performance, showing almost no evidence of learning the correct deductive conclusions via spontaneous thinking or further familiarization with binary inference problems. The experimental group with the multiple-choice problems shows significantly greater learning of the correct deductive conclusions with roughly 70 % of the subjects learning to draw correct conclusions to the first three binary inference problems and 50 % of the subjects learning to solve EC–NEC$_2$. Forcing a subject to use the analytic-deductive language for a series of multiple-choice problems produced substantial learning of the correct analytic-deductive conclusions to the binary inference problems as demonstrated on a subsequent constructed-response test, by subjects who were drawing incorrect conclusions on the test prior to the multiple-choice series.

(*d*) *Transfer from informational-analysis problems to multiple-choice problems*

The experimental group and the three informational-analysis control groups were all given the multiple-choice problems after the second test. Table 4 reports the performance on the third test, which followed the multiple-choice problems, for the subjects from the experimental and informational-analysis control groups who answered the binary inference problems incorrectly on the second test. Recall that the experimental group and the informational-analysis control groups received the same multiple-choice problems and differed only in that the experimental group had had prior training in the informational-analysis problems and the control groups had not.

Table 4. *Transfer from informational-analysis to multiple-choice problems*

(Number of subjects giving D conclusions on T3 of those subjects giving ND conclusions on T2.)

	2 EC$_1$			2 EC$_2$			EC–NEC$_1$			EC–NEC$_2$		
Group	D	ND	*P*	D	ND	*P*	D	ND	*P*	D	ND	*P*
Experimental	7	4	—	7	4	—	7	3	—	9	9	—
Informational-analysis spontaneous-thinking control	4	7	—	4	6	—	1	9	—	3	12	—
Informational-analysis problem-familiarization control	5	8	—	4	8	—	5	9	—	1	15	—
Informational-analysis concept-guessing-denial control	6	8	—	7	7	—	5	11	—	4	13	—
Total informational-analysis control	15	23	—	15	21	—	11	29	*	8	40	**

D, correct deductive conclusion; ND, not correct deductive conclusion. T2, second test of the binary inference problems; T3, third test. *P*, * $< 0{\cdot}05$, ** $< 0{\cdot}01$, one-tailed probability of difference from the experimental group (Fisher's Exact Test or χ^2).

The informational-analysis control groups are not significantly different from each other. The experimental and informational-analysis control groups are not significantly different on the 2EC problems, but the differences are in the expected direction. For the EC–NEC problems, the differences are significant. It should be noted that since a significantly higher percentage of subjects in the experimental group answered the problems correctly on the second test, there is heavy subject selection operating *against* the experimental group in later tests. Despite this, the remaining subjects in the experimental group show better learning in the multiple-choice problems than do the control groups. Thus, the results are even more significant than they appear.

Even among subjects who did not learn the correct analytic-deductive conclusions on the preceding test, learning of these conclusions in the multiple-choice problems is facilitated by prior experience with the informational-analysis problems. Since the subjects considered in this analysis had not learned the correct analytic-deductive conclusions to binary inference problems, the superior transfer of the experimental subjects must be due to something else that was learned from the informational-analysis problems. In the absence of any reasonable counter-conjecture, this provides some evidence that the informational-analysis cue evokes a general analytic-deductive method rather than directly evoking a concept attainment analytic-deductive method or specific concept attainment conclusions.

(e) *Transfer from inference to selection problems*

On the final test of the binary inference problems, given after eight binary inference problems *accompanied by feedback*, subjects in all six groups produced an average of 85 % correct analytic-deductive conclusions, but many more subjects in the informational-analysis and multiple-choice control groups than in the experimental group had learned the correct conclusions *only* when the answers were shown them. There were no significant differences between the six groups in inferences on the final test of the binary inference problems. Almost all of the subjects in each group had learned the correct analytic-deductive conclusions to the binary inference problems, but of course this does not mean that all had learned the same thing. To demonstrate that what is learned in the informational-analysis and multiple-choice problems is a general method (set, habit, tendency, etc.) of analytic-deduction in concept attainment problems, and not a particular set of verbal responses, a selection problem was given to all subjects after the final test of the inference problems.

The analytic-deductive selection method (focusing) involves varying one or a few dimensions in each selection to determine if the dimensions are relevant or irrelevant. The global-hypothesis selection method (scanning) involves varying all the dimensions except those included in the hypothesis about the concept. It is impossible to decide with perfect accuracy from selection data alone whether a subject is using the analytic-deductive method or the global-hypothesis method on any given trial. However, if all that is desired is a measure of the relative frequency of analytic-deductive selection methods, as opposed to global-hypothesis selection methods, in different groups on a given trial, then a statistic such as the number of subjects who vary only one dimension on that trial is a satisfactory index. The optimal selection method for the problem in this experiment was to vary only one dimension on a given trial. Subjects making analytic-deductive selections would vary fewer dimensions on the average than subjects making global-hypothesis selections, and the choice of a cut-off point, whether at one, two, or three varied dimensions on a trial, is a matter best determined by the maximally powerful statistical procedure of dividing at the median. For our experiments this meant a cut-off between one and two dimensions varied on a given trial.

The first trial of the selection problem is most sensitive to differences in transfer from prior experience because substantial learning to use the analytic-deductive selection method occurs in the course of a single selection problem. The three informational-analysis control groups were combined for the purposes of comparison with the experimental group because there were no significant differences among them in either inference or selection performance. The two multiple-choice control groups were combined for the same reason. Table 5 reports the number of subjects in the experimental, informational-analysis control, and multiple-choice control groups that made analytic-deductive selections (varied only one dimension from the initial example given to all subjects) on the first trial of the selection problem.

The experimental group, which received both the informational-analysis problems and the multiple-choice problems, made significantly more analytic-deductive selections on the first trial of the selection problem than the control groups, which lacked *either* informational-analysis problems *or* multiple-choice problems. Both the multiple-choice problems and the informational-analysis problems produce learning of an analytic-deductive method for concept attainment that transfers from inference to selection problems.

Table 5. *First trial selection method*

Group	Number of dimensions varied on first trial		P
	1	2–6	
Experimental	29	2	—
Multiple-choice control	24	12	**
Informational-analysis control	35	21	**

P, $** < 0{\cdot}01$, one-tailed probability of difference from experimental group (Fisher's Exact Test).

What is learned is a general method of solving concept attainment problems, not a particular set of verbal responses. Subjects in the control groups have learned the correct verbal responses to make in the inference problems, but this produces far less transfer to the selection problem than the general analytic-deductive method which is learned in the informational-analysis problems by some subjects and in the multiple-choice problems by other subjects.

IV. Conclusions

If we take the point of view that problem solving results from long chains of associations to the stimuli presented in a problem, then it may happen that an individual will fail to generate the correct solution to a problem not because his associative memory lacks the appropriate concepts for solving the problem, but because these concepts are not activated by the stimuli in the problem. If adequate stimuli for the activation of these concepts are added to the problem, in the form of either new information or new questions, then solution is achieved.

It bears mentioning that what an individual learns in the solution of any given problem may be a general method for solving a class of problems, not just the answer to the given problem. This experiment asked two *types* of questions about how individuals solve for the analytic-deductive conclusions to binary inference problems. First what stimuli evoke the analytic-deductive conclusions? Secondly, what is learned when an individual solves for the analytic-deductive conclusions to these problems?

It has been established that questions directing a subject to analyse information into two categories—possessed and needed—result in a substantial increase in the incidence of analytic-deductive conclusions in binary inference problems. This implies nothing about the importance of any other cues for evoking analytic-deductive conclusions, and, by itself, it implies nothing about the internal 'processing' of this cue that eventually results in analytic-deductive conclusions. However, this experiment has also established that the informational-analysis cue evokes an analytic-deductive *method* for concept attainment, not a specific set of conclusions to the binary inference problems. This was established by demonstrating that the informational-analysis training on binary inference problems transfers to a selection problem to produce a higher initial incidence of analytic-deductive selections. The responses required in selection problems are totally different from those required in inference problems.

Furthermore, we have also provided suggestive evidence that the informational-analysis cue produces learning of the analytic-deductive method for concept attainment by first evoking a general analytic-deductive method which operates on the binary inference problems to produce the analytic-deductive method for concept attainment. This hypothesis is supported by the superior learning of the analytic-deductive conclusions in the multiple-choice problems on the part of subjects who had had prior experience with the informational-analysis problems, but who had *failed* to learn the analytic-deductive conclusions for concept attainment in the informational-analysis problems. These subjects have clearly learned something, even though they have not learned the analytic-deductive method for concept attainment. This something facilitates learning to plan. Naturally we cannot be certain that this same something, which we have called a general analytic-deductive method, was elicited in the subjects who *learned* the analytic-deductive method for concept attainment in the informational-analysis problems, but this hypothesis appears much more likely as a result of the evidence obtained from subjects who *did not learn* the analytic-deductive method for concept attainment until the multiple-choice problems.

Perhaps it is not surprising that forcing a subject to use analytic-deductive conclusions in the multiple-choice problems results in a higher incidence of correct analytic-deductive conclusions in the binary inference problems, even after the subject is again free to use either global-hypothesis or analytic-deductive language. However, the positive transfer from the multiple-choice inference problems to the selection problem demonstrates that what is learned in the multiple-choice problems is not the specific conclusions alone, but a more general analytic-deductive method for concept attainment problems that results in analytic-deductive selections as well as analytic-deductive inferences.

Finally, it should be noted that all of the findings on the cues that evoke analytic-deductive methods, as well as the findings on what is learned when analytic-deductive methods are evoked by these cues, are findings valid for a substantial number of subjects, but certainly not all subjects. Inability to account for individual differences in some general theoretical framework is a defect of all studies of problem solving, but that makes it no less a defect.

REFERENCES

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

WICKELGREN, W. A. (1962). The planning method in concept attainment. Ph.D. thesis, University of California, Berkeley.

(*Manuscript received* 30 *May* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 155–163
Printed in Great Britain

SEMANTIC SATIATION OF WORDS AND NUMBERS*

By S. MESSER, L. A. JAKOBOVITS, R. KANUNGO and W. E. LAMBERT
McGill University

One purpose of the study reported is to compare the effect of continuous verbal repetition on the semantic ratings of words and numbers. A second is to examine the effectiveness of the semantic differential to index distinctive features of numbers as well as words. The procedure used is parallel to an earlier study on semantic satiation with words with the notable difference that, in the present case, a completely mechanized method is developed to present stimuli and record rating responses and their latencies. The following main findings are reported: (1) Regular semantic differential scales reflect significant satiation effects with words, but not with numbers, while a 'Meaningful-Meaningless' scale indexes significant shifts toward the 'Meaningless' pole with numbers, but not with words. (2) Certain differences are observed between words and numbers in the manner in which these are rated on standard semantic differential scales. (3) The M-shaped relationship between response latency and scale position reported earlier by Osgood is confirmed in the present paper. The implications of these findings in relation to the earlier study and with respect to the applicability of the semantic differential to numbers are discussed.

INTRODUCTION

The semantic differential is particularly useful as a method of determining distinctive characteristics of various types of stimuli. It has not only been frequently used as an indirect measure of the connotative meaning of words and word combinations (Osgood, Suci & Tannenbaum, 1957), but has also served as a convenient means of assessing the associational value of nonsense syllables (Fichman, 1959) and the distinctive aspects of non-verbal stimuli, including colour patches, pictures of objects (Osgood *et al.* 1957) and actual objects (Wimer & Lambert, 1959). Its versatility is evident, too, in its use as an index of changes in meaning resulting from the continuous verbal repetition of words (Lambert & Jakobovits, 1960) and the continuous observation of objects (Jakobovits & Lambert, 1964).

The commonly used form of the semantic differential was developed through factor analyses of subjects' ratings of common word stimuli on a wide range of semantic scales. One purpose of the present study was to examine its effectiveness as a method of differentiating the distinctive features of numbers as well as words. In a recent study, Jakobovits & Lambert (1962) argued that the meaning of numbers could be reduced through their continuous use, much as the meaning of words can be satiated (Lambert & Jakobovits, 1960). It was demonstrated that continuous verbal repetition of a number such as '7' prior to its occurrence in a simple addition problem ($7+4=?$) increased the time taken to add. Because the design called for rapid switching from repetition of a number to adding, no attempt was made to measure meaning changes of numbers with the semantic differential. The second purpose of the present experiment was to compare the effect of continuous verbal repetition on the semantic ratings of both words and numbers.

---

* This research was supported in part by the Canadian Defence Research Board, Grant D 77–9401–10, and in part by a subvention to W. E. Lambert from the Carnegie Corporation of New York.

METHOD

Subjects

Eighty McGill undergraduate students, none of whom had participated in an experiment of this nature, were tested invididually. The subjects were second-, third- and fourth-year undergraduates.

Material and apparatus

The symbols presented were six commonly used words and three numbers. The six words were divided into two sets of three each (Set X: *bus, city, smoke*; Set Y: *water, sky, key*). The numbers were either in digit form (1, 4, 9) or in word form (*one, four, nine*). Six control words (*lake hammer, tree, castle, car, bank*) and three control numbers both in digit form (2, 3, 8) and in word form (*two, three, eight*) were employed in the Different Word Control procedure, described below.

Each word and number was typed in capital letters on a slide to be projected on a ground-glass screen by means of a Kodak automatic slide projector. The seven points of the semantic differential scales were represented by seven electric keys (numbered 1 to 7 from left to right). These were connected in series to seven writing pens of an Esterline-Angus Operation Recorder (Model AW). The keys were $\frac{1}{2}$ in. wide and arranged in a straight line separated from each other by $1\frac{1}{2}$ in. intervals. They were placed on the table in front of the screen where the subject could comfortably reach them from his sitting position. Four semantic differential scales representing the three major factors were used in this experiment (*Good–Bad*; *Unpleasant–Pleasant*; *Hard–Soft*; *Passive–Active*). A fifth scale, *Meaningful–Meaningless*, similar to that used by Lambert & Jakobovits (1960), was added. The bipolar adjectives of each of the five scales were printed on slides in such a manner that the two adjectives appeared one under the other, with an arrow above the first adjective pointing left and an arrow below the second pointing right. The subjects were trained to use the keys on the left and right of centre to indicate how appropriate each adjective was for the stimulus under consideration.

The subject was instructed to put his index finger on a button placed near the centre of the bank of keys. After pressing one of the seven keys, he was to return his finger to the 'ready' button in preparation for his next response. The slide projector and the operation recorder, located out of view behind the ground glass screen, were connected to a time circuit in such a manner that each change of slide and each key pressing response were recorded on the chart of the Esterline-Angus apparatus. The time circuit was made to change a series of slides at certain fixed pre-arranged intervals. Since the projector was automatic, all the experimenter had to do was change the trays containing the slides. A two-minute rest period was given between the initial series and the final series while the experimenter changed the tape on the time circuit.

Procedure

Subjects were randomly assigned to one of four different groups, each composed of 20 subjects. One group performed under the experimental satiation condition and the other groups under three control conditions: *silence control*, *different-word control*, and *retest control*. In each group 10 subjects were assigned the words of set X and the numbers in digit form. To counterbalance the effect of order of presentation, five of these 10 subjects received the words before the numbers and the remaining 5 received them after the numbers. The other 10 subjects were assigned the words of set Y and the numbers in word form. The effect of order of presentation was controlled in the same manner as above.

All subjects were shown how to indicate their semantic ratings of words and numbers by means of the keys in front of them. Key positions were defined in terms of the scale positions of the semantic differential. They were given practice in making semantic ratings of figures, words, nonsense words, and numbers. They were frequently asked to report verbally what they meant by their responses, to ensure that they understood the instructions. No instruction about speed of responding was given. Subjects were simply told to act on their 'first impression' of the stimulus.

The first phase of the experiment proper was to obtain normal semantic profile ratings of the stimuli and latencies for making the ratings. Since there were three words and three numbers, each rated on five scales, 30 initial responses were recorded for each subject.

Satiation condition. After obtaining the initial response measures for the 20 subjects in this group, a word (or number) was exposed on the screen for 15 sec. and the subject was asked to repeat it aloud at a rate of two to three repetitions per sec. while looking at the word (or number) on the screen. After 15 sec. a scale was exposed and this subject made his rating by pressing one of the seven keys. The order of presentation maximized the separation of the recurrence of each word (or number) and each scale.

Silence control condition. The same procedure, with one exception, was used with 20 other subjects: they did not repeat aloud the words or numbers during the 15 sec. interval between the exposures of the stimulus symbol and the semantic scale. They were instructed to wait until a scale was presented before responding. Thus, a word (or number) was briefly exposed, then a 15 sec. interval elapsed before the semantic scale was shown, at which time the subject gave his rating.

Different word control condition. Using the same general procedure, another 20 subjects repeated aloud a particular control word or number during the time interval, and were then presented with a different word or number to rate. For example, the number '8' was exposed and the subject repeated it aloud for 15 sec. Then the number '1' was immediately exposed, and the subject was asked to pronounce and rate the number '1' on a semantic differential scale. This control condition was introduced to determine what effect the act of repeating aloud had upon the stability of the ratings. Whereas the previous condition is an 'unfilled-interval' control, the present is a 'filled-interval' control.

Retest control condition. The 20 subjects in this group merely repeated the initial series, i.e. they rated the same words and numbers twice with a rest interval of 2 min. between initial and final ratings.

Scoring

The semantic differential scores were transferred from the record chart. For latencies, each subject's record chart was analysed by means of a calibrated ruler which allowed for a maximum error of 0·08 sec. for individual measurements. Scoring reliability was checked by having a second judge, who was not aware of direction or meaning of differences, analyse a sample of 25 responses. The product-moment correlation coefficient between the two judges was found to be 0·99.

RESULTS

Three different response measures were analysed in the present experiment: the ratings on the four regular semantic differential scales; the ratings on the 'Meaningful–Meaningless' scale; and the latencies of rating responses. Each measure will be dealt with separately.

Regular semantic differential ratings

Semantic ratings for words. The means of the initial and final semantic scores for the three words on four scales were calculated for each of the four groups. These means represent polarity or extremity scores as used earlier by Lambert & Jakobovits (1960). The mean differences between these two scores, representing change of extremity of meaning, were also calculated and their significance levels determined. These are presented in Table 1. The Satiation and Different Word Control groups show significant amounts of meaning decrement following the repetition treatment and the other two control groups show no semantic change. A $4 \times 2 \times 2$ analysis of variance was performed on the mean difference scores to isolate the effects of the four different treatment conditions, the two sets of words (X and Y) used in each condition, and the order of presentation. The results reveal that there is a significant treatment-conditions effect with respect to meaning change ($F = 3{\cdot}45$, $P < 0{\cdot}05$). Neither of the two other main effects is significant ($F = 2{\cdot}49$ and $F = 1{\cdot}64$, respectively). The triple order interaction was similarly insignificant ($F = 1{\cdot}12$). The Satiation group is significantly different from the Silence Control group at the 0·01 level and from the Retest Control group at the 0·05 level, while the difference between the Satiation and the Different Word Control groups is not reliable. Furthermore, the three control groups do not differ significantly from each other.

Table 1. *Average change in polarity for three words or three numbers, summed over four scales*

		Change in polarity scores for									
		Words					Numbers				
Condition	No.	Initial mean	Final mean	Mean diff.	S.D. diff.	t	Initial mean	Final mean	Mean diff.	S.D. diff.	t
Satiation	20	21·20	17·05	−4·15	4·60	3·92**	10·65	9·35	−1·20	4·71	1·20
Silence control	20	21·80	22·00	+0·20	4·17	0·21	13·75	12·40	−1·35	4·05	1·45
Different word control	20	20·10	17·75	−2·35	4·26	2·40*	10·85	11·95	+1·10	3·01	1·59
Retest control	20	19·60	18·70	−0·90	4·15	0·95	13·50	12·25	−1·25	4·64	1·18

* $P < 0·05$. ** $P < 0·001$.

The entries are group means representing the sums of 12 ratings per subject, the maximum possible score being 36.

Semantic ratings for numbers. The mean difference between the initial and the final extremity ratings, representing change of intensity of meaning, was calculated for each group and none of these mean differences was found to be significantly different from zero (Table 1). A $4 \times 2 \times 2$ analysis of variance was performed for the numbers. Of the three main effects, only 'order of presentation' was significant ($F = 4·99$, $P < 0·05$). None of the interaction effects reached significance (the largest F value being 1·61). In order to examine the nature of the significant order effect, the means of the initial ratings were computed for the 40 subjects who were presented the numbers first, i.e. before the words, and for the other 40 subjects who were presented the numbers last, i.e. after the words. The means were 14·25 and 12·55, respectively, and these are significantly different ($t = 2·25$, $P < 0·05$).

Ratings on the Meaningful–Meaningless scale

The mean differences between the initial and final ratings on this scale were calculated for words and numbers and their significance levels were determined. These results are presented in Table 2. The continuously repeated numbers became significantly more meaningless whereas the change in the expected direction for repeated words did not reach acceptable significance levels (Table 2).

Latencies of semantic rating responses

An analysis of variance was performed for differences between the initial and the final latencies for both words and numbers. None of the main effects or interactions was significant, indicating that the Satiation group did not differ from the control groups with regard to response latencies. However, latencies generally decreased from the initial to the final ratings for most subjects. A correlated t test between the initial and final latency measures for all 80 subjects reveals a decrease significant at the 0·01 level (mean $= -1·66$, $t = 2·77$) in the case of the words, and at the 0·05

Table 2. *Average change in polarity for three words or three numbers on the Meaningful–Meaningless scale*

		Change in polarity			
		Words		Numbers	
Condition	No.	Mean diff.	S.D. diff.	Mean diff.	S.D. diff.
Satiation	20	−1·20	4·72	−2·45*	5·24
Silence control	20	−0·30	2·09	+0·10	3·19
Different word control	20	−0·50	2·73	−1·25	3·03
Retest control	20	+0·25	1·84	+0·25	3·33

* $P < 0{\cdot}05$.

A negative sign indicates a shift of subjects' rating response away from the 'meaningful' and towards the 'meaningless' end of the scale. A shift in the opposite direction is indicated by a plus sign.

level (mean $= -1{\cdot}55$, $t = 2{\cdot}25$) in the case of the numbers. This decrease in latency is presumed to be due to accumulated practice in responding to the scales.

Comparison of modes of rating words and numbers

In order to determine whether the standard scales of the semantic differential are appropriate for rating both numbers and words, detailed analyses of the ratings of both types of stimuli were made.

Fig. 1 presents the distribution of ratings of both words and numbers over the 7-point scale of the semantic differential. The ratings for all subjects ($N = 80$) were summed over all stimuli and all regular scales. The maximum possible frequency of usage of a point is 960 (80 subjects × 3 stimuli × 4 scales) and would occur if all subjects used that position for all their ratings. Fig. 1 reveals certain differences and similarities in the manner of rating words and numbers. First, the 'positive' half of the scales (that defined by *extremely*, *quite* or *slightly good*, *pleasant*, *hard* or *active*) is used more often than the opposite half (*bad*, *unpleasant*, *soft*, *passive*) both with numbers and words (using a sign test, $P < 0{\cdot}01$ in all four cases, i.e. initial and final ratings for both numbers and words). Secondly, the middle position, defined as 'neutral, irrelevant or meaningless', is used more often with numbers than with words ($P < 0{\cdot}01$ with sign test). This last observation is in agreement with the data given in Table 1 indicating that the means of initial extremity scores for words are markedly larger than the corresponding means for numbers. Thirdly, the positive points on the scales are used more often for words than for numbers ($P < 0{\cdot}01$ with sign test). However, there is no significant difference between words and numbers in the frequency of using negative scale positions, this being relatively low for both types of stimuli.

Fig. 2 presents the distribution of ratings over the 7-point Meaningful–Meaningless scale for both words and numbers. The ratings for all subjects ($N = 80$) were summed over all words or numbers. The maximum possible frequency of usage of a point is 240 (80 subjects × 3 stimuli) and would occur if all subjects used that position for all of their ratings. It will be noted: (1) that the positive points on the scale (Meaningful) are used more often than the negative positions (Meaningless); (2) the middle position

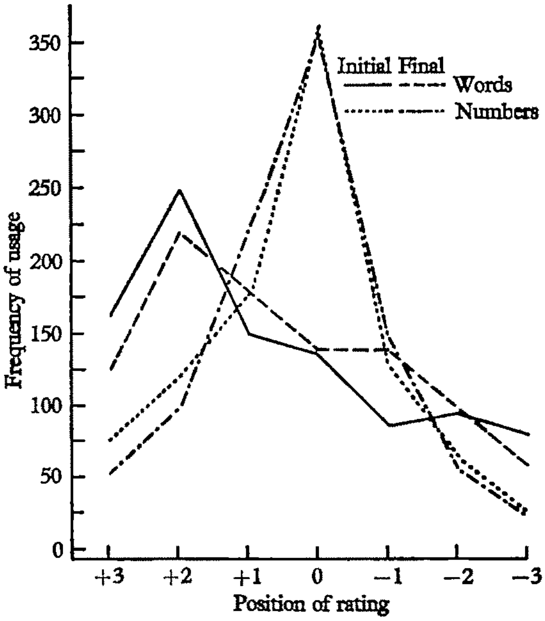


Fig. 1. Distribution of ratings of both words and numbers over the 7-point scale of the semantic differential. Each point represents the sum of ratings of 80 subjects over three stimuli and four scales.

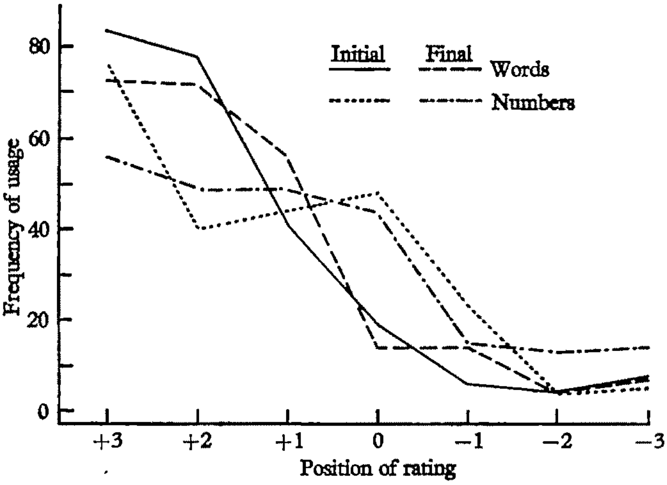


Fig. 2. Distribution of ratings of both words and numbers over the 7-point Meaningful–Meaningless scale. Each point represents the sum of ratings of 80 subjects over three stimuli.

is used more often with numbers than with words; (3) the positive points are used more frequently with words than with numbers. These three comparisons are in the same direction as those noted in Fig. 1 with regular semantic differential scales. However, the negative positions on this scale (Meaningless) are used significantly more often with numbers than with words. On this scale, too, as on the regular scales, the mean initial extremity of rating for the three words is significantly greater than for the numbers (for all four groups combined, the mean initial extremity of rating is 4·90 and 3·84 respectively, $t = 2{\cdot}29, P < 0{\cdot}05$).

Osgood (1941; also reported in Osgood *et al.* 1957, pp. 228–9) found shortest latencies for extreme judgements, longest latencies for intermediate judgements, with ratings in the middle position falling between these levels. Fig. 3 presents the distribution of latencies over the 7 points of regular semantic differential scales for both words and numbers. Fig. 3 and that presented by Osgood (1941) are very similar, even though different methods for measuring latencies were used (a voice key in the case of Osgood's data). Inspection of Fig. 3 reveals no systematic difference between words and numbers. It is interesting to note that the M-shaped relationship between response latencies and scale position holds for both initial and final ratings, irrespective of treatment and of the significant improvement effect noted earlier. It seems, then, that scale position is the determining factor of latency of response on the semantic differential for both numbers and words.

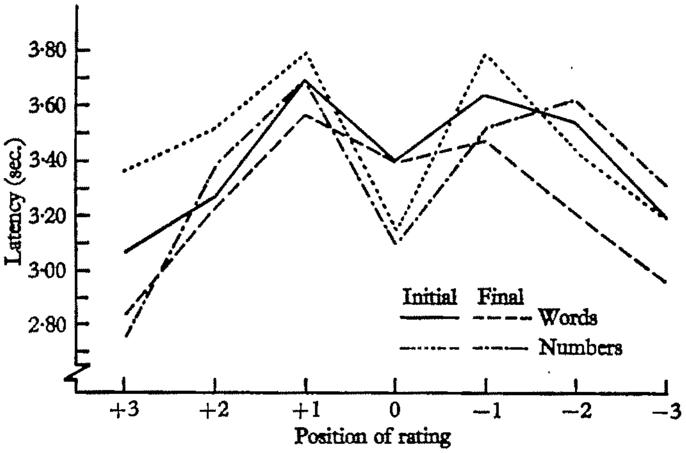


Fig. 3. Latency of response of standard semantic differential ratings as a function of their scale position. Each point represents the average latency per rating for 80 subjects.

DISCUSSION

The effects of continuous repetition on intensity of meaning. Continuous verbal repetition seems to have different effects on words and numbers when indexed with regular semantic differential scales. With words as stimuli, massed repetition induces a significant decrease in the intensity of semantic ratings made on standard semantic-differential scales. This finding is of special interest because it supports earlier work on satiation with an important change in procedure—the use of a fully mechanized method for presenting stimuli and recording both the position and latencies of ratings.

Two findings differ from those obtained in the earlier study of Lambert & Jakobovits (1960). In the 1960 study, words given satiation treatment moved towards the 'meaningless' and 'incomprehensible' poles of the scales 'meaningful–meaningless' and 'comprehensible–incomprehensible', and when the ratings for both scales were combined, the change was significant. The insignificant change toward the 'meaningless' pole in the present study may be due either to the failure to include the 'comprehensible' scale (making for less reliability) or to the fact that subjects in this case rated both numbers and words. A closer examination of such a possible interaction suggests that it may have a differential effect upon the reliability with which the Meaningful–Meaningless scale indexes satiation of words as compared to the regular semantic differential scales. Thus, the mean initial–final difference score on this scale for the 10 subjects in the Satiation condition who received order A (words first, numbers second) was $-3{\cdot}00$, indicating a change toward the Meaningless point, while the corresponding mean for the 10 subjects receiving order B (numbers first, words second) was 0·60, indicating a slight change toward the Meaningful point. However, this difference does not reach statistical significance ($t = 1{\cdot}75$, $0{\cdot}10 > P > 0{\cdot}05$). On the other hand, no such trend is apparent when the satiation effect is indexed by means of the regular semantic differential scales (the corresponding means are $-3{\cdot}50$ and $-4{\cdot}10$ respectively, $t = 0{\cdot}31$). Without the benefit of additional data, it is difficult to speculate on the nature of this interaction.

In the second place, the Different Word Control condition was not significantly different from the Satiation condition, although the difference was in the direction expected. In the 1960 study, the Different Word Control also yielded a large negative mean, as was the case here, but in the earlier study it was reliably different from the Satiation condition. Both studies are in agreement with the finding that the Different Word control condition does not differ significantly from the Silence and Retest control conditions. It is to be noted that in the Different Word control condition the subject was asked to repeat one word for 15 sec., then to repeat a different word one time, and finally to rate the second word on a semantic scale. This switch of attention from one word to another may be difficult to understand or manage even after full instructions and practice are given. There are two potential sources of error that can occur when using this procedure: the subject could inadvertently rate the first word—the one he has seen and repeated for 15 sec.—rather than the second, or he could rate some combination of the first and second words, and in either case he would be rating at least a partially satiated stimulus. Errors of this sort could account for the relatively large negative means noted for the Different Word Control condition, in both the present and the earlier study. It would seem, therefore, that it is not an appropriate control in comparison with the Satiation condition and other controls, none of which involves a rapid switch in stimuli to be rated. In cases where one is mainly interested in the effect of such a switch on encoding and decoding processes (as was the case in the bilingual study of Jakobovits & Lambert, 1962), the procedure is appropriate and particularly valuable.

In summary, the results indicate: (1) That continuous repetition of words significantly reduces the intensity of their meanings, as measured by the standard scales of the semantic differential. This change in meaning is attributable neither to the unreliability of ratings nor to the delay period before rating the second time, since both Retest and Silence controls showed no reliable changes. However, the meaning change is not reliably reflected on the Meaningful–Meaningless scale. (2) Because there were no differences between conditions in latency of rating, it appears that the satiation effect on words does not involve indecision about how to rate the word. Instead, subjects in the Satiation condition show no hesitation in making less extreme ratings following the repetition treatment when compared with controls whose ratings do not become less extreme. In fact, latencies in rating are determined more by the scale position being used than by variations in satiation or control procedures. (3) In the case of numbers, the satiation treatment did not affect the intensity of ratings on standard scales of the semantic differential but did induce a reliable shift in ratings toward the Meaningless pole of the Meaningful–Meaningless dimension. (4) The M-shaped relationship between response latencies and scale position, noted earlier by Osgood, holds for both numbers and words.

Semantic differential ratings of words and numbers. Words are rated much more frequently than numbers in the extreme positions (positive) of the semantic differential scales. Since extremity of rating is positively correlated with meaningfulness (Staats & Staats, 1959), numbers are judged as less meaningful than words by our subjects, a conclusion that is supported by the finding that words are rated as significantly more meaningful than numbers on the Meaningful–Meaningless scale. Furthermore, the middle scale position is used significantly more frequently for rating numbers than for rating words. These contrasting results suggest that the standard scales of the semantic differential, at least those used here, are not sensitive to the distinctive characteristics of numbers. This does not mean that other scales could not be chosen which would be more appropriate for numbers. In fact, the similarities noted in modes of rating words and numbers (apparent in both the latency data, Fig. 3, and in the tendency to use the positive scale positions more frequently than the negative, Fig. 1) suggest that a different form of the semantic differential would reveal the essential characteristics of numbers. The satiation treatment apparently does affect the connotative meaning of numbers since subjects indicated that numbers became reliably less meaningful after continuous verbal repetition. This finding lends support to the argument presented by Jakobovits & Lambert (1962) that the meaning of numbers is reduced by continuous verbal repetition making it more difficult to use them in addition problems.

REFERENCES

FICHMAN, S. H. (1959). The relation between association value of nonsense syllables and their meaning. *Undergraduate Research Projects in Psychology*, Vol. 1, Department of Psychology, McGill University. (Mimeo).

JAKOBOVITS, L. A. & LAMBERT, W. E. (1962). Semantic satiation in an addition task. *Canad. J. Psychol.* **16**, 112–19.

JAKOBOVITS, L. A. & LAMBERT, W. E. (1964). Stimulus-characteristics as determinants of semantic changes with repetition. *Amer. J. Psychol.* (in the Press).

LAMBERT, W. E. & JAKOBOVITS, L. A. (1960). Verbal satiation and changes in the intensity of meaning. *J. Exp. Psychol.* **60**, 376–83.

OSGOOD, C. E. (1941). Ease of individual judgement process in relation to polarization of attitudes in the culture. *J. Soc. Psychol.* **14**, 403–18.

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. (1957). *The Measurement of Meaning.* Urbana: University of Illinois Press.

STAATS, A. W. & STAATS, C. K. (1959). Meaning and *m*: separate but correlated. *Psychol. Rev.* **66**, 136–44.

WIMER, C. C. & LAMBERT, W. E. (1959). The differential effects of words and object stimuli on the learning of paired associates. *J. Exp. Psychol.* **57**, 31–6.

(*Manuscript received* 17 *June* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 165–180
Printed in Great Britain

THE EFFECTS OF CONTEXT ON THE VISUAL DURATION THRESHOLD FOR WORDS

By JOHN MORTON
*Department of Psychology, University of Reading**

Visual duration thresholds for words were measured under three conditions: (i) with a highly predictive context; (ii) with a lesser predictive context; (iii) a control condition with no context. The results show that the visual duration threshold for a word is reduced in the presence of a context by an amount depending on the probability of the word given the context. This result agrees with previous related findings.

The errors made by the subjects could be classified into three main groups. The first of these groups showed the effect of expectancy, or response bias. Further errors showed the influence of previous stimuli or responses and the third group indicated the role of the stimulus in determining the response. A model for the recognition of words is outlined which accounts for these results and attempts to define what is meant by 'the effects of response bias in the perception of words'.

I. Introduction

A hypothesis concerning the improvement of reading

It has been established (see Morton, 1959, 1961) that the reading efficiency of literate adults and students can be improved without recourse to the training of oculo-motor or 'perceptual' skills. Indeed, except in a few extreme cases, such training seems irrelevant to the problem. In addition, while it is apparent that 'intellectual skills' may be trained with profit, it is possible for a person to improve his reading proficiency merely by reading 'more actively' and concentrating more. This implies that there is some factor, or factors, relevant to the process of reading, which is not being fully utilized. The hypothesis is put forward that one of these factors is related to the statistical properties of language, or, more precisely, that the factor can be described in terms of such properties.

Thus we can consider that a reader, having language experience characteristic of the language as a whole, utilizes some of the available context in order to 'predict' the immediate stimulus. This would have the effect of making the immediate stimulus more probable, and the reader would require fewer visual cues to perceive it.

Assuming that such an effect takes place, then as a reader concentrates more in reading, he might be able to utilize the preceding context more effectively in predicting the immediate stimulus, facilitating its perception even more. It must be noted that the act of prediction is not envisaged as being in any sense deliberate. It is viewed merely as a function of the brain mechanisms involved in reading.

The effects of context in language behaviour

The hypothesis makes the assumption that context facilitates performance in verbal behaviour. Such an effect has been shown by several recent studies. 'Context' will be taken to include any factor which limits the number of possible stimuli, or increases the probability of occurrence of particular stimuli. In information theory terms, the stimulus contains less information, or is more redundant.

---

* Now at Applied Psychology Research Unit of the Medical Research Council, Cambridge.

In language studies there are two general methods of varying the information in a stimulus. The first is by utilizing the existing probability structure of the language, in terms of letters or words. To use the probability structure of the language in such a way, one does not assume that all individuals act at any instant as though they are aware of this structure but merely that there is a certain degree of communality among the subjects used to generate the material and those tested with it, with respect to their experience of language.

The second method involves the experimenter artificially imposing constraints upon the experimental situation, for example, by varying the number of words in a list from which the response has to be made.

Studies in the psychology of language have used both these methods of varying the information in a stimulus. The results show that increasing redundancy affects the recall of words (Miller & Selfridge, 1950), nonsense syllables (Rubenstein & Aborn, 1954) and letters (G. A. Miller, 1958) and the learning of words (Sharp, 1958), nonsense syllables (Baddeley, 1960 personal communication) and letters (Adelson, Muckler & Williams, 1955).

The effects of context in recognition tasks have been shown for both visual and auditory presentation with a variety of contexts.

Miller, Heise & Lichten (1951) and O'Neill (1957) both found that words presented auditorally were more readily recognized in noise when presented in a sentence than in isolation. Both experiments showed that the removal of the context raised the threshold for 50 % correct recognition by 6 db. Miller *et al.* (1951) also showed that with monosyllables the threshold was lowered by presenting the subject with a list of words from which the stimulus was drawn. The ease of recognition was inversely related to the number of items in the list. Under similar conditions, I. Miller (1957) found that nonsense bi-syllables were more readily recognized in redundant lists. Bruce (1958) compared the auditory recognition of sentences in isolation, in the context of a correct description of their topic, and with an incorrect description of their topic. He found that the correct context led to an increase in the number of words recognized, and the false context reduced the score.

In a visual recognition task, Miller, Bruner & Postman (1954), exposing 8-letter sequences of zero-, 1st-, 2nd- and 4th-order approximations for times ranging from 10 to 500 msec., found that the number of letters recognized increased with the order of approximation over the whole range of exposure times. E. E. Miller (1956) presented two groups of 25 sentences either in random order or in an order where the sequence told a story, and found that significantly more words were perceived in the 'story' condition.

No work has been found on the effect of a sentence context on the visual perception of words, which is the paradigm for reading. It is necessary that context should operate in this situation if the original hypothesis is true. We would expect the presentation of a sentence context to facilitate the perception of a word. In addition, from concepts derived from information theory, we would expect there to be a linear relationship between the ease of perception of the word and the logarithm of the probability of it occurring given the context. The ease of perception of the stimulus words are measured in the present experiment by measuring their visual duration thresholds.

II. MATERIALS

It was necessary to prepare a series of words with context which predicted them to a more or less specifiable extent. The contexts were incomplete sentences which could be completed by the addition of a single word; e.g.

It was a most exciting ——.
She looked into the ——.
He walked round the ——.

The incomplete sentences are referred to hereafter simply as 'sentences'.

It seemed possible that the properties of the sentence, apart from the stimulus, might exert an influence on the threshold. That is, different sentences might set up different preparatory states in the subject, depending on the number and nature of the responses possible.

To investigate this point it seemed advisable to have more than one word associated with each sentence. The method adopted for doing this was to have one word with a high transitional probability and one with a low transitional probability connected with each sentence, and to have each word connected in a like way with two sentences. In this way the sentence effect, if any, could be examined. 'High' was defined arbitrarily as $P \geqslant 0{\cdot}15$, 'low' as $P < 0{\cdot}15$. The two classes were termed Condition 1 and Condition 2 respectively.

Accordingly, 30 sentences were designed which seemed likely to produce responses from which a satisfactory selection of words and sentences could be made. The sentences were arranged in lists such that there was maximum possible separation between those sentences likely to produce a similar response, in order to reduce interference effects. They were then presented one at a time to subjects with the instruction to respond immediately by saying the first word which occurred to them which completed the sentence grammatically and logically. If any response was outside these conditions the sentence was presented again. In this way a hundred responses to each sentence were collected from a hundred female students.

Table 1. *Examples of response configurations*

1. Sentence: The cup was placed on the —.

Responses:

table	67	drainer	1
saucer	16	chair	1
sideboard	6	chest	1
shelf	4	cupboard	1
mantlepiece	3		

2. Sentence: They went to see the new ——.

film	22	baby	2
house	20	bedroom	1
show	14	hospital	1
play	12	exhibition	1
car	6	bungalow	1
building	6	curtains	1
picture	3	calf	1
moon	3	child	1
flat	2	toys	1
shop	2		

The responses for each sentence were summed for each word, the frequencies of occurrence of the words being used as an estimate of their transitional probabilities. Examples of the configurations obtained are given in Table 1. The lists of words and sentences were examined for combinations fulfilling the conditions stated above. Since the sequences were incomplete, a further 25 sentences were prepared, and the same procedure was followed with another 100 female students. Among the 55 sentences there were 24 sentences and words which could be suitably combined, of which 18 were used in the experiment. These 18 sentences and words are given in Table 2. It will be seen that the word 'coat', for example, occurs as a high probability word in the context of sentence 1, and as a low probability word in the context of sentence 3.

To control for possible differences in threshold due to the configuration of the word in isolation, a third condition was used, employing a row of 'x's' instead of a context sentence, to give the same stimulus conditions.

The three conditions of presentation were then:

Condition 1—high probability context.
Condition 2—low probability context.
Condition 3—control condition, no context.

The sentences and words were printed in indian ink on white card in lower case, using U.N.O. stencils, and photographic slides were prepared.

Table 2. *The stimulus words and context sentences*

	Words	
Sentences	Condition 1	Condition 2
(1) Coming in he took off his	coat: 32	shoes: 14
(2) She cleaned the dirt from her	shoes: 44	dress: 6
(3) She bought a new	dress: 32	coat: 10
(4) Together they had the first	dance: 17	prize: 2
(5) He hoped to win the	prize: 24	cup: 4
(6) At the sink she washed a	cup: 27	saucer 2
(7) The cup was placed on the	saucer: 16	shelf: 4
(8) She reached up to dust the	shelf: 22	window: 3
(9) She gazed thoughtfully through the	window: 89	door: 4
(10) Quickly she walked towards the	door: 44	shop: 3
(11) He bought a book at the	shop: 70	station: 4
(12) He asked the way to the	station: 36	cinema: 6
(13) They went to the	cinema: 21	dance: 4
(14) They went to see the new	film: 22	picture: 3
(15) They looked intently at the	picture: 29	film: 4
(16) He was wakened by the	birds: 16	noise: 13
(17) She was startled by the	noise: 45	train: 2
(18) In the distance she heard the	train: 16	birds: 4

The figure following each word is the number of times the word was given as a response to the sentence in 100 subjects.

III. APPARATUS

The apparatus consisted of a rigid white cardboard screen, two projectors and a falling plate tachistoscope. The screen contained a slit of transparent paper on which the context sentence, or the row of 'x's', was projected. The beam of the other lantern passed through the tachistoscope opening to permit the word to be exposed for a limited amount of time, at the end of the sentence and precisely aligned with it. The lens of the second lantern was stopped to give a sharp on-set and cut-off of illumination as the tachistoscope plates fell, and the intensities of the beams were adjusted to give equal brightness on the screen. The subject sat facing the screen at a distance of 4 ft. in a slightly darkened room. The letters on the screen were a quarter of an inch high and could be read clearly by all the subjects. The tachistoscope was checked with a photoelectric cell and a Chronotron timer and was found to have a consistency of ± 3% over the whole range.

IV. PROCEDURE

The words were arranged in an order which gave maximum separation between those connected with the same sentence, and three lists were drawn up, as far as possible at random, assigning contexts to the words. Constraints were applied so that each list contained six of each of the three conditions. The three lists of words and conditions may be seen in Table 3.

It was not possible to present any word more than once to any subject, since previous work indicates that the threshold for a word would be affected by a recent presentation. Hence each subject was presented with only one list and saw each word only once in one of the three conditions.

Since previous work indicated that there was a marked practice effect in the tachistoscopic recognition of words (e.g. Howes & Solomon, 1951), two further precautions were taken. First, at least three practice words were given, one paired with each of the three conditions. The second precaution was to present the words in four alternative orders: 1–18; 18–1; 9–1; 18–10; 10–18; 1–9. Thus there were three lists with four possible orders of presentation each.

The subjects were 24 female students who were assigned at random to one of three groups of eight, each group being given one of the three lists. Two subjects in each group had the list in each of the four orders. Any subjects with defective vision wore their normal correcting lenses.

The subjects were tested singly. The following instructions were given: 'This is an experiment to find out how we perceive words. I'm going to flash words on to the screen here, and I want you to tell me what word or part of a word you see or think you see. I don't want you to guess though. The words will be presented in two ways, either with an incomplete sentence in front of it, which the word completes grammatically and logically, or with a row of 'x's' in front. In both cases your procedure will be the same. I will not necessarily tell you whether your response is correct or not.'

The practice words were then presented and the subjects became familiar with the procedure. For each run the sentence was projected on to the screen until the run was completed. After the subject had read the sentence the signal 'ready' was given, and the word exposed. The method of ascending limits was used, the exposure time being raised by increments of 11·54 msec (equivalent to two units of the arbitrary scale on the tachistoscope) to a criterion of two successive correct responses. Following one correct response, the exposure time was kept the same for the next exposure. The starting-point for each run varied with the subject, word and condition, and was based on an estimate such that three incorrect, partial or 'nil' responses might be given before the correct one. All partial or incorrect responses were noted.

After the experiment, without previous warning, each subject was asked to recall all the words she had seen. When no more words could be recalled, those sentences were read out comprising the contexts for the words not remembered and additional responses were noted.

Table 3. *Mean threshold for words (in milliseconds)*

	Group I ($N = 8$)		Group II ($N = 8$)		Group III ($N = 8$)	
Word	Condition	Threshold	Condition	Threshold	Condition	Threshold
(1) coat	3	107	1	75	2	81
(2) dance	3	117	2	96	1	97
(3) saucer	1	76	2	74	3	104
(4) door	2	86	3	124	1	86
(5) cinema	1	90	3	102	2	89
(6) birds	1	70	2	92	3	104
(7) shoes	2	74	3	119	1	97
(8) prize	2	89	1	96	3	86
(9) shelf	3	99	1	70	2	79
(10) shop	3	86	2	76	1	56
(11) film	2	104	1	96	3	128
(12) noise	3	119	1	93	2	94
(13) dress	1	75	2	109	3	139
(14) cup	2	99	3	119	1	81
(15) window	2	70	3	100	1	56
(16) station	1	54	2	73	3	115
(17) picture	3	80	1	84	2	86
(18) train	1	59	3	99	2	81
Mean for the group		85·9		94·3		92·1

V. RESULTS

(*a*) *Threshold*

The raw results are summarized in Table 3.

Since the order effects were negligible compared with the individual variation, the results were pooled and submitted to an analysis of variance. The analysis is complicated owing to the fact that only a third of the subject-word-condition combinations occur, and is described elsewhere (Morton, 1961). The results are summarized in Table 4.

Table 4. *Analysis of variance*

Source	D.F.	S.S.	M.S.*†	*F*	*P*
(1) Groups	2	164·70	82·35[1]	0·581	N.S.
(2) Subjects within groups	21	2978·26	141·82[2]	9·0	< 0·001
(3) Conditions: 1 *v.* 2	1	121·42	121·42[4]	4·56	< 0·05
(4) Conditions: 1 and 2 *v.* 3	1	1965·06	1965·06[4]	73·81	< 0·001
(5) Conditions × groups	4	107·76	26·94[3]	1·086	N.S.
(6) Conditions × subjects/groups	42	734·65	17·49[2]	1·116	N.S.
(7) Words	17	1459·18	85·83[3]	3·461	< 0·01
(8) Residual (words)	28	694·39	24·80[2]	1·583	< 0·05
(9) Residual (groups)	315	4937·44	15·67	—	—
	413	13162·86			

* Mean squares tested against: [1] Subjects within groups; [2] Residual (groups); [3] Residual (words); [4] Special error term equal to ($\text{M.S.}_8 + \text{M.S.}_6 - \text{M.S.}_9$) (see Morton, 1961) $= 24{\cdot}80 + 17{\cdot}49 - 15{\cdot}67 = 26{\cdot}62$.

† The error term for the conditions × groups source should also contain a component from the main words variance, and from the conditions × subject within group variance, but, since this source is not significant in any case, it is unnecessary to include these terms.

The following conclusions may be drawn from the analysis (the numbering of the conclusions corresponds to the numbering of the sources in the table):

(1) The difference between the groups is no more than would be expected from the differences between subjects.

(2) Subjects vary significantly in what might be called their basic threshold. This may be related to the degree of certainty required before a decision is made as to the nature of a stimulus (this as an attribute of the discrimination processes in the brain; cf. Tanner & Swets, 1954).

(3) and (4) As predicted, the threshold under Condition 1 is significantly lower than under Condition 2, and both context conditions have a very significant effect compared with the control condition.

(5) There is no difference between the groups in their effective use of context. This result, together with conclusion (1), means that we can pool the data from all three groups for further examination.

(6) Within each group of eight subjects there were significant (though only just significant) differences between the individuals in their effective use of context.

(7) The stimulus words which were used differed basically in their ease of perception. This is discussed below.

(8) The residual (word) term contains some variance due to the word × condition interaction which we would expect to be significant, since within the context conditions there is a considerable variation as to the actual transitional probability.

The results were further analysed in various ways. In these analyses the assumption was made that the uncontrolled variables affecting the threshold, i.e. group differences and word differences, did so in an additive way. Accordingly, the mean thresholds of each word in each condition were adjusted in the following way.

The group mean thresholds across all words and conditions in the arbitrary units were: group I, 9·14; group II, 10·31; group III, 10·00. To allow for individual differences in basic threshold, figures of 1·17 and 0·86 were then subtracted from the mean scores per word of group II and group III respectively, to make the three group means equal. To correct for the basic differences in the thresholds of the words, the mean score of a word in Condition C (control), corrected for group differences, was subtracted from the mean corrected scores for that word in Conditions 1 and 2, and a constant was added to make all the resulting figures positive. All comparisons of thresholds in the context conditions thus use figures which have the thresholds of the words under the control condition as a base-line. In the figures presented, these adjusted threshold measures are treated as being in arbitrary units.

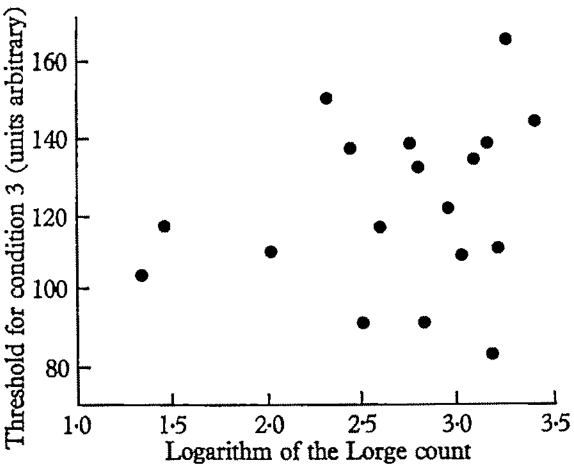


Fig. 1. The relationship between threshold and frequency of word occurrence in the absence of context.

(*b*) *Word differences*

According to Howes & Solomon (see Howes & Solomon, 1951; Solomon & Howes, 1951; Howes, 1954) the visual duration threshold for words has a negative linear relationship with the logarithm of the probability of the words in the language. However, when the thresholds for the control condition, corrected for group differences, were plotted against the logarithm of the Lorge count (Thorndike & Lorge, 1944), no such relationship could be found (Fig. 1). All the words are common ones though, and it is unlikely that effects of differential experience of words, the factor used to explain the result, could operate appreciably within such a small frequency range. There was no relation between threshold and such structural factors as length of word or the number of ascenders and descenders. However, a factor of 'confusability' of the words gave a rank-order correlation with threshold in Condition 3 of $\tau = 0{\cdot}456$ ($P < 0{\cdot}015$). The factor of confusability was estimated by taking the number of different whole words given as incorrect responses under the control conditions.

(*c*) *The effect of context*

The analysis of variance showed a significant difference between the two experimental conditions. In order to examine more closely the relation between threshold and probability, it seemed reasonable to use the results only of those subjects who did exhibit an effect due to context. Thus the four subjects from each group with the largest mean context effects were selected, and their results were treated in the following analysis.

Fig. 2 shows a plot of the mean threshold for the word against the logarithm of the probability ($\log P$). The thresholds have been corrected for group differences and word differences, so they are an estimate of the relative effects of abstracted context-stimulus situations. Each plot is the mean for four subjects. The product-moment correlation between $\log P$ and this threshold measure was $r = 0{\cdot}528$ ($P < 0{\cdot}01$).

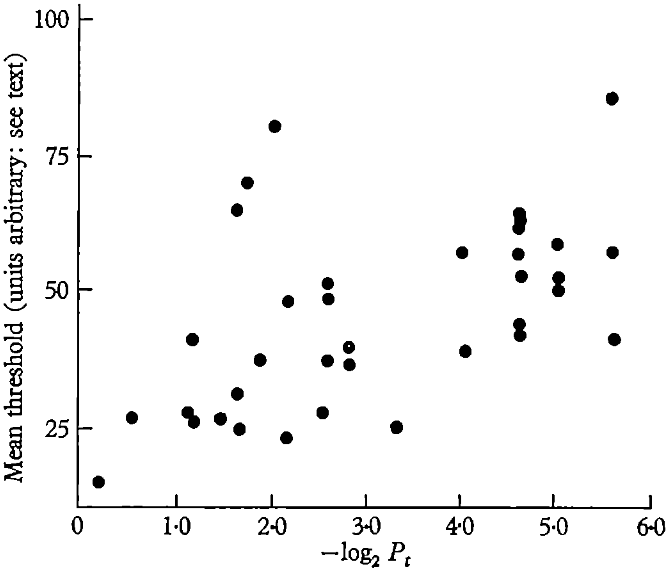


Fig. 2. The relationship between threshold and transitional probability of a word in a sentence context.

The experiment was also designed to separate out the effect of the sentences upon the thresholds. However, no significant relationship could be found between the thresholds and the entropy of the context sentences (estimated by $H = \Sigma - p \log P$ over all possible responses to the sentence).

(*d*) *Memory*

The test of memory for the stimulus words, given at the end of the experiment, produced the following results. (i) There was a steady increase in recall with serial order. (ii) Significantly more words were recalled in the context conditions than in the control condition. Ease of recall of the words was highly determined by the relative ease of recall of the context sentences.

(*e*) *Errors*

The errors made by the subjects were analysed in order to provide evidence as to the relative importance of the stimulus and of response bias. The incorrect responses fell into three groups showing the effects of expectancy, of previous stimuli and responses, and of inadequate perception. The first two of these groups do, in one sense, reveal the effects of response bias, in that the influences operating in determining the response are not predominantly those relating to the physical nature of the stimulus. We will, however, account for these errors in terms other than 'response bias' in the 'guessing' sense used by Goldiamond & Hawkins (1958).

Table 5. *Examples of incorrect responses involving a word with a higher transitional probability* (*see also Table* 6)

Condition 2

Stimulus	Response	Stimulus	Response
coat	dress, hat (2)	birds	bells
dance	cinema, theatre (2)	shoes	coat
door	window	shelf	table (3)
cinema	station	film	book

Condition 1

Stimulus	Response
coat	hat
saucer	table (2)
prize	race (3), —ce (2)

Note. A figure following a word indicates the number of times that word was given as a response.

The effect of expectancy

In Conditions 1 and 2, there was a bias towards giving as an incorrect response a word more probable in the context than the stimulus word, compared with a less probable word. In Condition 1 there were seven higher probability words and thirteen lower probability words given as incorrect responses. In Condition 2 the figures were 23 and 32 for higher and lower probability words respectively. All higher probability words given as incorrect responses are shown in Table 5. As there are considerably more lower probability words available as responses in the experimental conditions, it seems as though an expectancy for a more probable word exerts an influence on perception, resulting in a higher proportion of such words being given as incorrect responses than would be expected. Each subject who gave a higher probability word as a response was questioned after the experiment as to the confidence with which this response was made. In most cases they reported that they had been completely certain, and in some cases suspected that the stimulus word had been changed in the middle of a run. It seems clear that the result cannot be explained by saying that the subjects were 'guessing'.

Thirty-five words were given as responses in the context conditions which were nonsense in the context. If the subjects' responses were considered to be guesses determined by the context, they would not have made these responses. This point is considered further in the discussion which follows.

Influence of previous stimuli and responses

Another class of error involving the influence of some kind of set consists of those cases where the response was influenced by a word previously seen in the experiment. These responses, words and parts of words, are shown in Table 6. From the table it is clear that a previously seen, higher probability word is particularly prone to be given as an incorrect response on the next trial. None of the stimulus words which had not yet been seen were given as errors, except in the five cases where they were more probable words. So we can say that the effect is genuinely due to the prior presentation of the word with which the subject responded.

Table 6. *Examples of the influence of a previous stimulus word*

	Conditions 1 and 2		Condition 3	
	Stimulus	Response*	Stimulus	Response
Immediately preceding stimulus word	door	cinema	coat	train
	shoes	prize	dance	saucer
	film	show (shop)	saucer	dance
		----p (shop)	shop	shelf
		---se (noise)	shoes	prize
	noise	fl-- (film)	birds	c---n- (cinema)
Earlier stimulus word	shoes	dress (×2, P 6)†‡	birds	dress (×2, P 7)
		c--- (coat, P 6, H)§	prize	dress P 6)
	dress	coat (P 12)	dress	fl-- (film, P 2)
	coat	dress (P 12, H)	cup	shop (P 4)
	dance	cinema (P 3, H)	picture	-ance (dance, P 3)
	cinema	station (P 12, H)		
	birds	noise (P 6, H)		
	prize	dance (P 6, H)		

* The response word was the previous stimulus word except where a word is given in parentheses. † (P *n*), the word occurred *n* words previously. ‡ (×2), such a response was made by two subjects. § H, a higher probability word.

A similar influence was the effect of a previous incorrect response. One example was that of the subject presented with 'train' in Condition 3 who responded 'sting'. On the next exposure the incorrect initial 's' was retained and the response was 'strain'. Similarly, with a stimulus of 'shelf', the response 'smell' was followed by 'shell'; and 'tree' and 'true' in response to 'cup' led to 'trip' and 'trap' as further responses before the word was reported correctly.

Perceptual errors

The errors can also be classified in terms of what appear to be various kinds of perceptual error. In many cases an error may be classified in more than one category, and may also have been classified above. Examples of these errors are given in Tables 7 (i) to (v). The five categories in Table 7 are considered in paragraphs (i) to (v) below.

(i) The most common error, totalling 176, involves the correct perception of beginning and/or the end of the word, while the remainder of the word is incorrectly reported. In over half these errors the response word was the same length as the stimulus word.

(ii) Ascending and descending letters were confused on 138 occasions. Thus 't', the most frequent of the ascending letters, was mistaken for another ascending letter 17 times, and reported as being seen when absent on 44 occasions. No reliable conclusions may be drawn as to the effects of letter frequency in this situation, however, since there was no control of the number of occurrences of the different letters, or of the various other influences mentioned.

The last three categories of error are examples of what might be called 'stimulus disorientation'. It would appear that the shape of the word or the form of a letter have been perceived correctly, but an error has been made in either the horizontal or the vertical orientation of a generalized shape category.

Table 7. *Examples of perceptual errors*

	Stimulus	Response		Stimulus	Response
(i) Incorrect centres	coat	coot, soot, roat	(iii) Inversions of ascending and descending letters	dance	pence, grace, gence, p---ance, p----
	door	dear			
	birds	bards, bands, beads		shop	prop
	shop	sharp		dress	grass
	cup	cap, crap		picture	lecture, structure
	picture	puncture		train	grain
(ii) Confusions of ascending and descending letters	cost	coal	(iv) Inversions of w, n and u	window	windem
	dance	trance, trace, tence		train	treasu, ----u
	door	bear, boar		dance	---use, --u--
	shelf	shell, shield		picture	--tn-
	film	tile (2), table			
	dress	stress, class	(v) Word reversals	coat	t---(2), to--
	train	brain		shop	p---

(iii) On 32 occasions, ascending letters were seen as descending letters, in 12 cases the reverse was found. The most common inversion was 'd' seen as 'g' its exact inversion in the print used, apart from the curl in the tail.

(iv) Other inversions involved 'w' seen as 'm' and confusions of 'u' and 'n'.

(v) In 32 cases it seemed as if a word had been reversed. It was judged so when a word with an initial ascending or descending letter was seen as having a final ascending or descending letter, or vice versa.

The justification for regarding the last three categories as cases of stimulus disorientation lies in the unambiguous examples such as 'train' reported as ending in 'u' and 'coat' as beginning with 't'. The phenomenon of stimulus inversion was most strikingly demonstrated by one subject who was presented with the word 'dress' in the control condition. Her responses alternated between 'dress' and 'grass' five times before she settled for the correct response. She commented that she had had no idea which way up the initial letter was.

The types of errors reported here are substantially the same as those noted by previous workers such as Pillsbury (1897) and Vernon (1929). Although the conditions of perception in the present experiment are widely different from those experienced in reading, Morton (1961, 1964) has shown that the same kinds of error occur in both situations.

VI. Discussion

The minimal role of response rise in the experiment

A relationship has been shown between the transitional probability of a stimulus and its visual duration threshold. This effect would be predicted on the basis of the previous experiments showing the effects of context on perception which were cited above. The validity of ascribing some of these effects to perception has been challenged by Goldiamond & Hawkins (1958). Their subjects were first given a training session in which they repeated nonsense syllables at different frequencies. They were then told, in a recognition session, that these words would be flashed subliminally; they were to guess until accurate. Accuracy was defined as responding with the word on the experimenter's score sheet. Blank exposures were presented throughout. The experiment mimics an ascending Method of Limits with 'accuracy' terminating a series, and a negative linear relationship was found between the logarithm of frequency of training and 'threshold'. The authors regard their results as challenging a perceptual interpretation of the word-frequency-threshold relationship, suggesting that 'response bias' could explain the results adequately. (Note: for the moment, the traditional perceptual-response dichotomy is accepted without comment.)

Since their subjects were specifically instructed to guess, their use of the term 'response bias' is a very special one. The following points may be made in asserting that this kind of response bias played a small part in the present experiment.

(i) The subjects were specifically instructed not to guess.

(ii) The 'accuracy' criterion in the present experiment was for two successive correct responses, and on several occasions the subjects reported uncertainty as to the nature of the stimulus on the exposure following a correct response.

(iii) The mean transitional probabilities of the words in Conditions 1 and 2 were 0·334 and 0·049 respectively. Following the Goldiamond and Hawkins result we would expect the proportion of correct to incorrect initial responses in the two conditions to be approximately 35:65 and 5:95, i.e. 1:1·8 and 1:19, respectively, if a response bias determined the response. In fact, these proportions were only 1:0·45 for Condition 1 and 1:0·65 for Condition 2.

(iv) The subjects often reported that the words which were given as incorrect responses were seen quite clearly. Only occasionally was a response prefaced by 'I think it was...'; and the response which followed was as often correct as incorrect.

The use of visual cues in determining the response

If the effect of context upon threshold is accepted as a perceptual phenomenon, then we can state that in the context conditions a subject will perceive a word at an exposure time which would be inadequate for the perception of that word in the absence of a context. In other words, the presence of a context reduces the number of visual cues necessary for the correct identification of the word.

By considering that a high proportion of the errors result from a response being made on the basis of insufficient visual cues, it seems possible, from the nature of these errors, to say something about the order of emergence of stimulus detail with exposure time, i.e. to rank those stimulus properties which seem to determine a response. Two factors are considered.

(*a*) A large number of the errors involved the confusion of ascending or descending letters. In addition, a high proportion of error words were the same length as the stimulus word. We can consider the two properties of configuration, that is the position of ascending and descending letters, and the length of the word, as specifying the *shape* of the stimulus. In terms of the figure-ground relationship, this is the 'contour' (cf. Vernon, 1952, p. 43).

(*b*) The largest category of error consisted of correctly reporting the beginning or the end of a word, while incorrectly identifying the interior letters. It may be noted that Woodworth & Schlosberg (1954, p. 104) refer to the phenomenon of the relative masking of the interior letters in a sequence. They refer to the effect as occurring in peripheral vision only, but their hypothesis that the effect is due to a differential brightness contrast could be applied equally well to foveal vision with very short exposures.

Certain of the errors were analysed in order to test whether the properties of shape and the relative ease with which outside letters are perceived had any special significance. These errors consisted of the first incorrect whole word responses in any run in the Control Condition, excluding responses of a previous stimulus word. In this way, as many errors as possible were excluded which could be attributed to the influence of set. The 87 responses which were thus chosen were analysed as follows:

(*a*) Nineteen responses had the correct configuration, 22 the correct length and a further 28 were correct in both respects. Thus a total of 69 of the 87 errors were correct with respect to some aspect of the shape.

(*b*) Twenty had the correct initial letter, 33 the correct final letter, and a further 12 had both letters correct—a total of 65 of the 87 responses.

(*c*) Only eight of the responses had none of these elements in common with the stimulus word.

When the 159 whole word errors in the Context Conditions are analysed, we find 88 with some aspect of the configuration matching the stimulus and 80 with either or both of the initial or final letter correct. Forty-nine of the responses had none of these factors in common with the stimulus, though of these about half were classified as involving stimulus inversion or reversal; the rest were mainly previous responses or highly probable words in the context.

To establish the significance of these findings would require a control experiment of some complexity in design, but it remains reasonable to suggest that these properties have a certain priority in determining the response.

It seems advisable to restate the present position.

(1) It was suggested that in the presence of context, fewer visual cues are required for a word to be perceived and reported correctly.

(2) It has been demonstrated that the errors made cannot, in the main, be attributed to the operation of a response bias, and must therefore be due to perceptual effects. This accepts the traditional perceptual-response dichotomy.

(3) It was suggested that the errors result from responses being made on the basis of insufficient visual cues, and certain properties of the stimulus have been identified which might be so utilized.

The effect of set

We now return to the main result of the present experiment and consider it further. It was established that context has the effect of reducing the visual duration threshold of a word, and that the amount of this reduction is related to the transitional probability of the stimulus word. Now the measures of transitional probability were established by presenting 100 subjects with the incomplete sentences and summing their responses. This we will term a *Generation Situation*. In assuming a communality of response probability for the population, we are stating that the transitional probability values measure the relative strengths of the response biases (in the sense used by Goldiamond & Hawkins) of any subject, given the context.

However, we have established that the responses in the experimental situation, which we will call a *Recognition Situation*, are not appreciably determined by response bias. Hence, we can state that the processes which operate to determine the response bias of a word in a context also operate to produce a reduction in threshold of the word in that context. These processes are concerned with the sequential relationships between linguistic units, and the nature of their operation presumably derives from an individual's experience of language. These processes will be called the *Sequential Processes*.

The result of the operation of these processes is the same in both the generation and recognition situations—the response of a word.

Let us now differentiate between those subprocesses in the brain which operate terminally to make different words available as responses, and call them *units*. These 'units' are similar to those suggested by Treisman (1960) as the components in some 'dictionary' of known words. The units may be regarded as having a threshold of activation; when the threshold is exceeded, the unit fires, and the corresponding word is available as a response. Normally, different units have different levels of activation, determined, for example, by their frequency of usage. (We may alternatively regard them as having different thresholds; that is, at the present time, purely a matter of taste—the principle of operation remains the same.)

Since the presence of a context makes certain words more likely to be given as responses, we can say that the Sequential Processes operate to increase selectively the level of activation in certain units.

In a *Generation Situation*, a selection mechanism operates to choose a word in such a way that the relative probability of different words being chosen depends upon the levels of activation in the corresponding units. From such responses we estimate the transitional probabilities, i.e. in Goldiamond & Hawkins's terms, the response bias.

In a *Recognition Situation*, we suggest, sensory information is fed into the set of units, or dictionary, as it becomes available. This input increases the level of activation of particular units until a unit fires. The word is then available as a response. The increase in activation of any unit will be a function of the degree of similarity between the stimulus and the word associated with the unit; that is, the number of 'visual cues' they have in common. In the presence of a sensory input into the dictionary the subjects report the available word as having been 'perceived'.

In this system, if a word is very highly probable in a context, then its unit will be at a high level of activation. Therefore minimal activation from the sensory input would be likely to trigger the unit, resulting in a lower threshold for the higher probability words and giving rise to incorrect responses of the first group mentioned.

The effects of previous responses may be predicted by considering that the threshold-frequency effect has been obtained with nonsense syllables by differential training immediately before measuring the thresholds (Solomon & Postman, 1952; Baker & Feldman, 1956; King-Ellison & Jenkins, 1954). Such training involved seeing and saying the syllables at different frequencies, which, from the differential effect on threshold, may be said to have differentially increased the activation level in the different units. We can then accept that perceiving and responding with a word in the present experiment will also temporarily raise the activation level in the appropriate unit. Thus the unit will be more likely to fire when sensory information is fed in, giving rise to incorrect responses of the second class. It might also be noted that similar errors occurred in a reading task where a class of intrusive errors was found consisting of words which had previously occurred in the passage, usually on the same line (Morton, 1961, 1963). This form of activation may be contrasted with the more permanent effects attributed to differential experience of different words in the language.

It was noted above, that high probability words which had previously occurred were particularly prone to be given as incorrect responses. In terms of the model this would indicate that the various activation effects are cumulative.

The third group of errors, perceptual errors, would be expected from the suggested mode of operation of the sensory input. The greater the similarity between a word and the stimulus word, the more the appropriate unit will be activated and the more likely will the word be given as a response.

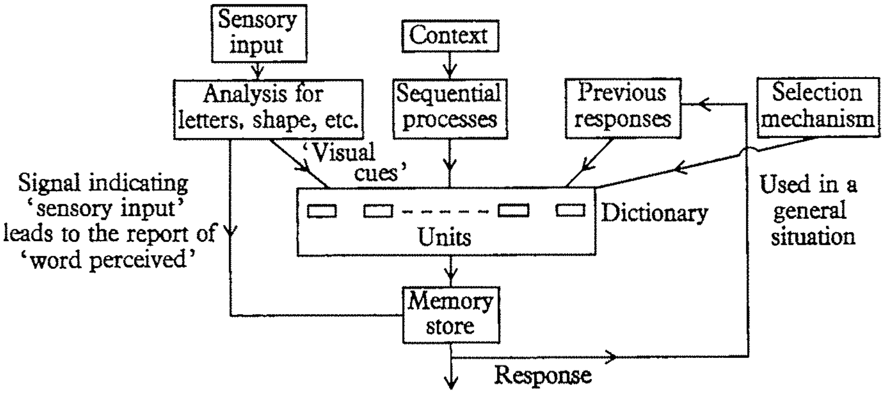

Sensory input
Context
Analysis for letters, shape, etc.
Sequential processes
Previous responses
Selection mechanism
'Visual cues'
Signal indicating 'sensory input' leads to the report of 'word perceived'
Units
Dictionary
Used in a general situation
Memory store
Response

Fig. 3. A model for word recognition.

The model is presented in the form of a flow diagram in Fig. 3. Its main virtue in its present form is that it eliminates the difficulties implicit in an arbitrary linear separation of stimulus and response effects. In the model, both effects take place at the same level, and the 'perception' of a word is only different from the generation of the same word in that there is an input from the sensory mechanisms in the first case, and an input from some Selection Mechanism in the second.

In its present form, the model is sufficient to account for the main results of the experiment and the errors made by the subjects. It is developed more formally elsewhere (Morton, 1961).

The work here described was performed at Reading University under the supervision of Prof. M. D. Vernon while the author was in receipt of a D.S.I.R. studentship. The author is also grateful to Dr E. R. F. W. Crossman and Dr E. C. Poulton for discussions at various stages in the work, and to Dr B. Wetherill of Birkbeck College, London, for producing the Analysis of Variance.

REFERENCES

ADELSON, M., MUCKLER, F. A. & WILLIAMS, A. C. (1955). Verbal learning and message variables related to amount of information. In N. Quastler (ed.), *Information Theory in Psychology*. Glencoe, Illinois.

BAKER, K. E. & FELDMAN, H. (1956). Threshold-luminance for recognition in relation to frequency of prior exposure. *Amer. J. Psychol.* **68**, 278–80.

BRUCE, D. J. (1958). The effect of listeners' anticipations on the intelligibility of heard speech. *Language and Speech*, **1**, 79–97.

CHERRY, C. (1957). *On Human Communication*. London: Chapman and Hall; New York: John Wiley.

GOLDIAMOND, I. & HAWKINS, W. F. (1958). Vexierversuch: the logarithmic relationship between word-frequency and recognition obtained in the absence of stimulus words. *J. Exp. Psychol.* **56**, 457–63.

HOWES, D. H. (1954). On the interpretation of word frequency as a variable affecting speed of recognition. *J. Exp. Psychol.* **48**, 106–12.

HOWES, D. H. & SOLOMON, R. L. (1951). Visual duration threshold as a function of word-probability. *J. Exp. Psychol.* **41**, 401–10.

KING-ELLISON, P. & JENKINS, J. J. (1954). The duration threshold of visual recognition as a function of word-frequency. *Amer. J. Psychol.* **67**, 700–3.

MILLER, E. E. (1956). Context in the perception of sentences. *Amer. J. Psychol.* **69**, 653–4.

MILLER, G. A. (1951). *Language and Communication*. New York: McGraw-Hill.

MILLER, G. A. (1958). Free recall of redundant strings of letters. *J. Exp. Psychol.* **56**, 485–91.

MILLER, G. A. & SELFRIDGE, J. A. (1950). Verbal context and the recall of meaningful material. *Amer. J. Psychol.* **63**, 176–85.

MILLER, G. A., BRUNER, J. S. & POSTMAN, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *J. Gen. Psychol.* **50**, 129–39.

MILLER, G. A., HEISE, G. A. & LICHTEN, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J. Exp. Psychol.* **41**, 329–35.

MILLER, I. (1957). Perception of nonsense passages in relation to amount of information and speech-to-noise ratio. *J. Exp. Psychol.* **53**, 388–93.

MORTON, J. (1959). An investigation into the effect of an adult reading efficiency course. *Occup. Psychol.* **33**, 222–37.

MORTON, J. (1961). Reading context and the perception of words. Unpublished Ph.D. thesis, Reading University.

MORTON, J. (1964). A model for continuous language behaviour. *Lang. Speech* (in press).

O'NEILL, J. J. (1957). Recognition of intelligibility test materials in context and isolation. *J. Speech Dis.* **22**, 87–90.

PILLSBURY, W. B. (1897). The reading of words: a study in apperception. *Amer. J. Psychol.* **8**, 315–93.

RUBENSTEIN, H. & ABORN, M. (1954). Immediate recall as a function of degree of organization and length of study period. *J. Exp. Psychol.* **48**, 146–52.

SHARP, H. C. (1958). The effect of contextual constraint upon recall of verbal passages. *Amer. J. Psychol.* **71**, 568–72.

SOLOMON, R. L. & HOWES, D. H. (1951). Word frequency, personal values and visual duration thresholds. *Psychol. Rev.* **58**, 256–70.

SOLOMON, R. L. & POSTMAN, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *J. Exp. Psychol.* **43**, 195–201.

TANNER, W. P. & SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychol. Rev.* **61**, 401–9.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York.

TREISMAN, A. (1960). Contextual cues in selective listening. *Quart. J. Exp. Psychol.* **12**, 242–8.

VERNON, M. D. (1929). The errors made in reading. *Spec. Rep. Ser. Med. Res. Counc.* no. 130.

VERNON, M. D. (1952). *A Further Study of Visual Perception*. Cambridge University Press.

WOODWORTH, R. S. & SCHLOSBERG, H. (1954). *Experimental Psychology*. New York: Holt.

(*Manuscript received* 11 *April* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 181–187
Printed in Great Britain

IMPRINTING AND SHORT-TERM RETENTION

By W. SLUCKIN and K. F. TAYLOR*
Department of Psychology, University of Leicester

In the first of three experiments chicks imprinted to a moving object were found to discriminate between the original and new moving objects significantly better after an interval of some minutes than immediately after 'training'. In the second experiment similarly imprinted chicks were tested for discrimination between the initial object and a strange one when these objects were stationary. In these circumstances discrimination tended to be rather poorer after a time interval than immediately after training. In the third experiment imprinted chicks were tested with the familiar object only for following, and not for discrimination, at different times after training. Following was found to improve after some separation from the object. This recovery from 'drive satiation' is thought to account for the reminiscence-like phenomenon found in the first experiment.

I. Introduction

The presence and strength of imprinting may be assessed in experimental studies by two criteria. Initially, the young animal, usually a newly hatched bird of a nidifugous species, is 'trained' with some figure, commonly a moving object, without any extraneous reinforcement. Then on a later occasion, its responses are either tested: (1) by confronting it again with the figure, or (2) in a choice situation, when the original figure and a different figure are presented simultaneously (Sluckin & Salzen, 1961). The former test is one of recognition, and the latter is one of discrimination.

Before imprinting is inferred from positive responses to the original figure in these tests, results obtained from control experiments must be considered. The recognition test of imprinting calls for a comparison of responses to the figure between 'trained' and 'untrained' subjects; imprinting is presumed when the approach and following responses of the experimental animals are significantly better than those of the 'untrained' control animals. The discrimination test of imprinting requires additional evidence concerning the inherent attractiveness of the figures between which the subject makes a choice; if the animals on test discriminate significantly in favour of that figure which was originally presented rather than in favour of one particular figure, then this is taken as evidence of imprinting. Most recent studies of imprinting have depended on the discrimination criterion. In the present study we are concerned with imprinting assessed in two experiments by discrimination criteria, and in a third experiment by the recognition criterion.

As days go by an attachment to an object is steadily 'forgotten'; cf. Salzen & Sluckin (1959), Guiton (1961). The present study is concerned with manifestations of imprinting some minutes, rather than some days, after training. In particular, we wished to compare the subjects' ability to discriminate between the original and new figures (1) under conditions resembling training, and (2) under conditions differing markedly from those at training.

---

* Now at Department of Psychology, University of Melbourne, Australia.

II. EXPERIMENT I. DISCRIMINATION BETWEEN MOVING OBJECTS

(*a*) *Subjects and apparatus.* The subjects were 48 chicks hatched in the laboratory from commercial eggs. Upon hatching, the neonates were transferred to individual cardboard pens. Light, heat, food and water were provided, but each subject was in visual isolation from the others. Experimental trials were conducted in a runway with walls of corrugated cardboard fixed to an angle-iron frame, with internal dimensions 10 ft. × 2 ft. 2 in. × 10 in. high. The linoleum flooring was divided by pencilled lines into ten sections, each 1 ft. × 2 ft. 2 in. Illumination and heating were provided by 150 W infra-red lamps, and food and water were kept in small containers placed half-way along the runway against one of the walls. A pulley system extended over the centre of the runway, enabling a stimulus object to be suspended some 2 in. above the floor and moved its entire length in both directions. An electric motor operated the pulley so that the object moved 1 ft. in 3 sec., then remained stationary for 3 sec., and so forth. Two stimulus objects were used: a red, black and white cardboard box, 5 in. long × 2·5 in. × 2·5 in.; and a sheet of green foam rubber folded into approximately pyramidal shape, 4 in. long × 1·5 in. × 4 in. high.

(*b*) *Procedure.* The behaviour of each subject was observed on the second day after the day of hatching, so that post-hatch age at 'training' and 'testing' ranged between 36 and 54 hr. Each subject was transported in a straw-lined box from its pen to the runway, and placed on the centre of the half-way line. The stimulus object was then set in motion from one end of the runway, the box and the foam being used in alternate trials for this purpose. Subject was left in the runway for a 'training' period of 3 hr., after which the object was stopped and removed for one of four intervals which preceded 'testing'. Twelve subjects were tested almost immediately after training and twelve after 5, 15 and 45 min. Under each of these four conditions half the chicks remained in the runway for the interval and half were placed in isolation in an adjacent pen. At the start of the test period the two stimulus objects were set in motion from opposite ends of the runway, and for 5 min. a commentary on the subject's behaviour was recorded on tape. The test was evaluated by playing back the commentary and analysing the subject's behaviour, on a time-scale, into the following categories:

Positive reactions: (*a*) orientate towards object; (*b*) approach object; (*c*) follow object; (*d*) contact object.

Negative reactions: (*a*) look away from object; (*b*) distress chirps on approach of object; (*c*) move or run away from object.

The observer's summary at the conclusion of the test was also noted. This summary, and the pattern of behaviour shown on the time-scale, enabled E to assign one of four scores to the trial:

−1 Clear preference for the new object.
0 No discrimination.
1 Moderate discrimination (some positive reaction to the original object OR some adverse reaction to the new object OR a combination of these for part of the trial).
2 Clear discrimination (strong positive reaction to original OR strong negative reaction to new OR a combination of these throughout the trial).

Each test was scored twice, with experimental conditions unknown to E; on the second occasion scores were made without knowledge of previous results. The few discrepancies were resolved by a third 'blind' scoring.

(*c*) *Results.* The discrimination scores of the 48 subjects in the first experiment were initially examined for any effect that could be attributed to the chicks spending the interval between training and testing in or out of the runway. As shown in Table 1, the differences between these two conditions were found to be small at all time-intervals; for *no* interval was P less than 0·93 by the two-tailed Kolmogorov–Smirnov test (Massey, 1951).

Table 1. *Discrimination by chicks after various intervals*

Duration of interval (min.)	Individual scores when interval spent in runway	Totals	Individual scores when interval spent out of runway	Totals
0	−1 0 1 0 2 0	2	1 1 0 0 0 1	3
5	1 0 2 2 1 0	6	0 2 0 0 0 1	3
15	2 2 0 1 1 0	6	2 1 1 2 0 2	8
45	2 1 2 2 1 1	9	1 2 1 1 1 0	6
Totals	—	23	—	20

Each entry in the table represents the discrimination score of one subject, the scores ranging from −1 to +2.

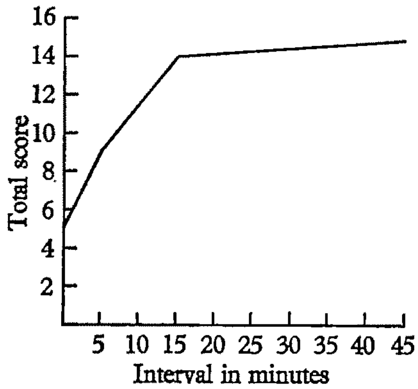


Fig. 1. Total discrimination scores after different intervals in Expt. I.

Fig. 1 sets out the total discrimination scores of the 48 subjects as a function of the duration of the interval between training and testing. The graph shows that discrimination was more evident as the interval lengthened. In fact, without any interval between training and testing, seven of the twelve birds failed to discriminate in favour of the familiar object, but this number fell to three after 15 min., and to only one out of twelve birds after 45 min. The upward trend of the curve in Fig. 1 is noteworthy; the difference in scores between the conditions of no interval and one of 45 min. is statistically significant by the two-tailed Mann–Whitney U-test ($P = 0{\cdot}02$).

III. Experiment II. Discrimination between stationary objects

(*a*) *Subjects and apparatus.* Subjects were 32 chicks, reared in the same conditions as in the first experiment. Two runways were used, the one already described and another of identical proportions.

(*b*) *Procedure.* Again each subject was carried in a box from its pen to a runway, and there trained with a moving object for 3 hr. The object was then stopped and removed either momentarily or for 45 min. Sixteen subjects were assigned to each condition. Immediately before testing, the subject was placed under a cardboard lid on the halfway line of the runway. The two objects were then suspended on the pulley at distances of one foot from the opposite ends of the runway. Both ends of the runway were employed an equal number of times for original and new objects. Each chick was observed during a 5 min. test period. It was noted (*a*) if subject approached to within 1 ft. of, and (*b*) if subject made physical contact with, either stationary object.

(*c*) *Results.* Table 2 shows the number of subjects in the second experiment that discriminated in favour of the stationary original object by approach and by contact after no interval and after 45 min. This time discrimination was observed rather more frequently after no interval than after the 45 min. interval. However, the difference between these two conditions was not significant statistically ($0{\cdot}2 < P < 0{\cdot}3$) either by the criterion of approach to within 1 ft. of the object or by the criterion of contact.

Table 2. *Number of chicks choosing a stationary original object*

	Interval	
Type of discrimination	0	45
Approach and contact	11/16	8/16
Approach only	13/16	9/16
None	3/16	7/16

It is interesting to note that only one of the 32 chicks used in this experiment actually approached the new object, and then only for 14 sec. The other 9 non-discriminators remained in the neutral area of the runway throughout the test period. Discrimination in favour of the original object rather than the new object was significant both at no interval (Fisher exact-probability test, one-tailed $P = 0{\cdot}025$) and after 45 min. (one-tailed $P < 0{\cdot}05$).

IV. EXPERIMENT III. FOLLOWING A 'FAMILIAR' OBJECT

(*a*) *Subjects and apparatus.* Thirty chicks were reared as in the first and second experiments, and the trials were conducted in the two runways already described.

(*b*) *Procedure.* Each subject was trained with a moving object for 3 hr., as in the previous experiments. Three interval conditions were then used, 10 subjects being assigned to each:

(1) No interval—object stopped at end of runway (alternated) for 30 sec., then re-started. Subject remained in runway (Control condition).

(2) 'Traumatic' interval—object stopped at end of runway (alternated) for 30 sec., then restarted. Subject remained in runway but was placed under cardboard lid for the interval.

(3) Thirty-minute interval—object stopped at end of runway (alternated) and removed for 30 min. Subject remained in runway.

After the interval the subject was tested with the object moving for 5 min. Following was recorded in two ways:

(1) Distance—the number of lines on runway floor crossed by subject while moving towards the object.

(2) Time—the number of seconds subject orientated and moving towards, or in same section of runway as, the object.

(c) *Results.* It can be seen from Table 3 that, by both criteria, the following of the subjects in the two experimental groups was clearly better than the following of the control group. Employing Mann–Whitney U-tests, the differences between the distance scores of the control group and the two experimental groups just failed to reach the 5 % level of significance. The differences between the time scores, however, were both significant (control *v.* trauma, two-tailed $P = 0{\cdot}02$; control *v.* interval, two-tailed $0{\cdot}002 < P < 0{\cdot}02$).

Table 3. *Following by chicks in various testing conditions*

	Distance. Mean percentage of possible lines crossed	Time. Mean no. of sec. spent following
Ten experimental birds—tested after 30 min. interval	62	134
Ten control birds—tested at finish of training	35	74
Ten experimental birds—tested after 'trauma'	60	136

V. DISCUSSION

In our first experiment, when both the old and the new objects moved during the test, immediately after training 5 out of 12 experimental chicks chose the old object and one chose the new one. However, 45 min. after training, under otherwise identical conditions, 11 out of 12 chicks chose the old object and none chose the new one. A delayed improvement in the retention of what has been learned is, in human verbal learning, termed 'reminiscence'. As the performance of our imprinted chicks was better after an interval than straight after training, we may for convenience say that our animals showed 'reminiscence'.

The occurrence of 'reminiscence' in these circumstances is not altogether surprising. An acquired attachment to a figure may be expected to show itself in an animal or a human being not so much after a momentary separation from the figure as after a separation lasting some time. Superior discrimination after an interval of 45 min. rather than immediately after training occurred, however, only in Expt. I and not in Expt. II: that is, only when the objects were moving during the test. Two problems present themselves, viz. (1) why 'reminiscence' occurs in imprinting at all, and (2) what factors suppress it.

The first problem may be considered in terms of the drive-satiation hypothesis mentioned by Guiton (1959). This postulates that the tendency to follow a figure decreases as the young animal is stimulated by the presence of others of its own kind. An extension of this hypothesis would be that the tendency to follow a figure also decreases as the animal continues to be stimulated by the figure itself, and increases again after this stimulation has ceased. Previously, it had been reported by Sluckin & Salzen (1961) that chicks followed a moving object less persistently after having spent some hours with it. In the present study we set out in Expt. III to investigate whether there would be any improvement in the following of a figure after a period of separation from it. We found that following scores were significantly higher after a 30 min. interval than they were immediately after training. Now, since an animal's ability to discriminate between objects cannot show itself to the full unless it follows well, much of any observed improvement in discrimination could be due to improved following. Thus, the rise in the graph for Expt. I may well have resulted from the chicks' recovery from drive-satiation.

Why, then, was there no recovery from drive-satiation after the longer interval in the second experiment? First it may be asked whether this had anything to do with the 'shock' sustained by the chicks when momentarily covered by the lid. In Expt. III this shock experience produced recovery from drive-satiation as effectively as did a period of separation from the figure; the shock should, therefore, have *facilitated* discrimination in Expt. II. Further, in the light of the finding in Expt. III just mentioned, it may also seem rather surprising that in Expt. I another 'trauma', removal from the runway, did not materially affect discrimination. It may be that the clear discrimination reached in this experiment after 15 min. or so could not be further improved by 'traumatic' experiences. However, a difference in discrimination might have been expected between those birds that had been in, and those that had been out of, the runway at 'no interval'. Comparisons between groups of only six birds may well be inconclusive. A full-scale experiment could be conducted to find out whether *discrimination* would improve after a 'trauma' just as *following* has been shown to improve.

In Expt. II there was no improvement in discrimination, and possibly some deterioration, after a 45 min. interval between training and testing with the objects stationary. Why this was so is at present open to speculation. It may be that recognition of an object remains unimpaired for short periods only when it has the same attributes, including motion, which it had during training. When recognition remains unimpaired, discrimination improves because recovery from drive-satiation produces better following. On the other hand, recognition may deteriorate rapidly as minutes between training and testing pass if an object lacks some of its original attributes. In such circumstances, recovery from drive-satiation cannot offset the loss of retention.

We do not know the relative importance of the different attributes of a figure for the retention of attachment to it. The contrast between the results of Expt. I and Expt. II suggests, at first sight, that movement is of crucial importance, but, contrary to the earlier view, it has been shown that imprinting does not necessarily depend on movement or visual flicker (Gray, 1960). What we have observed is that when *all* attributes of the training figure are retained during testing, a reminiscence-like rise in the retention curve occurs. Therefore, the question arises whether a similar result would be found if the object was presented without movement during *both* training and testing. Further experimental studies are needed to investigate the character of short-term retention of imprinting to stationary objects.

REFERENCES

GRAY, P. H. (1960). Evidence that retinal flicker is not a necessary condition of imprinting. *Science*, **132**, 1834.

GUITON, P. (1959). Socialisation and imprinting in Brown Leghorn chicks. *Anim. Behav.* **7**, 26–34.

GUITON, P. (1961). The influence of imprinting on the agonistic and courtship responses of the Brown Leghorn cock. *Anim. Behav.* **9**, 167–77.

MASSEY, F. J., Jr. (1951). The distribution of the maximum deviation between two sample cumulative step functions. *Ann. Math. Stat.* **22**, 125–8.

SALZEN, E. A. & SLUCKIN, W. (1959). The incidence of the following response and the duration of responsiveness in domestic fowl. *Anim. Behav.* **7**, 172–9.

SLUCKIN, W. & SALZEN, E. A. (1961). Imprinting and perceptual learning. *Quart. J. Exper. Psychol.* **13**, 65–77.

(*Manuscript received* 23 *April* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 189–199
Printed in Great Britain

FURTHER TESTS WITH THE SPIRAL AFTER-EFFECT OF A THEORY OF HOMOEOSTATIC EXCITATION–INHIBITION

BY C. G. COSTELLO
Psychological Laboratory, Regina General Hospital, Saskatchewan

On the basis of a theory of homoeostatic excitation–inhibition, a number of predictions concerning the spiral after-effect were tested in four experiments. The predicted effect of the interaction of direction of rotation of the spiral and duration of stimulation was confirmed. The predicted effect of the interaction of the above two variables: (1) with the extroversion–introversion variables tested in Expt. II, and (2) with stimulant *v.* depressant drugs tested in Expt. III, received only partial confirmation. The predicted alternating decrease and increase in the after-effect with massed practice and the effects of the interaction of massed practice with the direction of rotation were confirmed. The predicted effect of depressant and stimulant drugs interacting with massed practice and duration of rotation was not confirmed.

The results are discussed in relation to the usefulness of the theory of homoeostatic excitation–inhibition, and the value of molar neuropsychological theories is questioned.

INTRODUCTION

Previous work has indicated that the expansion spiral after-effect produced by clockwise rotation of the spiral is longer than the contraction after-effect produced by counter-clockwise rotation (Costello, 1960; Eysenck, Willett & Slater, 1962). There is also some evidence that massed practice and brain damage produce different effects on the two types of spiral after-effect (Gallese, 1956; Spivack & Levine, 1957; Holland & Beech, 1958; Philbrick, 1959; Costello, 1961).

A theory of homoeostatic excitation–inhibition has been proposed to account for these differences between the expansion and contraction after-effect (Costello, 1961). The experiments to be reported in this paper were designed to provide tests of predictions from the theory.

EXPERIMENT I

In the original presentation of the theory (Costello, 1961) an attempt was made along the following lines to account for the finding that the expansion after-effect is longer than the contraction after-effect. It was suggested that, during the clockwise rotation used to produce the expansion after-effect, most of the excitation produced could be expected to remain within the contour of the edge of the spiral disk. Following Deutsch (1956) it was assumed that a ‘wave front’ is propagated during spiral rotation in the direction of the motion of the contour of the spiral. During clockwise rotation this motion is towards the centre of the disk. A further assumption of isomorphism is of course made by such an explanation. It was suggested also that during anti-clockwise rotation some of the excitation spreads out beyond the contour produced by the edge of the spiral disk, since the direction of motion of the spiral figure is outward from the centre. Thus, after an equal period of stimulation, a greater amount of excitation will have accumulated within the contour of the spiral disk after clockwise than after counter-clockwise rotation. If we assume that the duration of the after-effect is proportional to the amount of excitation built up by stimulation, then it would follow that the expansion after-effect would be longer than the contraction after-effect.

Since the theory does not suggest that there is any limit to the amount of spread of excitation beyond the spiral contour during the counter-clockwise rotation, the following prediction follows: with an increase in duration of stimulation, the increase in the expansion after-effect (produced by clockwise rotation) will be greater than the increase in the contraction after-effect (produced by counter-clockwise rotation). The first experiment was designed to test this prediction.

Method

The spiral used had four throws of 180 degrees and rotated at 100 rev./min. It was 8½ in. in diameter and mounted 6 ft. from the observer.

The subjects used were 28 females and 12 males whose ages ranged from 17 to 29 years. They were allotted randomly to one of four groups.

All the subjects were given a practice 15 sec. stimulation with the spiral. After a rest period of 15 min., the subject was presented with one of the four following stimulating conditions: (1) 30 sec. clockwise rotation; (2) 30 sec. counter-clockwise rotation; (3) 120 sec. clockwise rotation; (4) 120 sec. counter-clockwise rotation.

Results

The mean duration of the after-effect for the four groups was as follows:

(1) 30 sec. clockwise rotation (expansion after-effect): 24·97 sec.
(2) 120 sec. clockwise rotation: 51·78 sec.
(3) 30 sec. counter-clockwise rotation (contraction after-effect): 18·64 sec.
(4) 120 sec. counter-clockwise rotation: 19·30 sec.

An analysis of variance of the data (Table 1) indicated: (1) That both the main effects of direction of rotation and duration of stimulation were significant. (2) The interaction between direction of rotation and duration of stimulation was significant.

Table 1. *Analysis of variance of the data obtained on direction of rotation and duration of stimulation (Expt. I)*

Source	Sum of squares	D.F.	M.S.V.	*P*
Due to differences in direction	3765·54	1	3765·54	< 0·001
Due to differences in duration of stimulation	1886·77	1	1886·77	< 0·005
Due to direction : duration interaction	1707·47	1	1707·47	< 0·01
Residual error	6705·81	36	186·27	
Total	14065·59	39		

In line with the results of previous work, the expansion after-effect was longer than the contraction after-effect. In line with prediction from the theory of homoeostatic excitation–inhibition, there was a significant increase in duration of the expansion after-effect from 30 to 120 sec. stimulation and a slight non-significant increase in the duration of the contraction after-effect from 30 to 120 sec. stimulation.

EXPERIMENT II

Previous work has suggested that experimental conditions designed to produce cortical inhibition in the sense used by Eysenck (Eysenck, 1957) result in a shortening of the duration of the spiral after-effect. Thus administration of depressant drugs has been found to reduce the duration of the after-effect (Eysenck, Holland & Trouton, 1957; Costello, 1963) and a similar reduction has been produced by massed practice (Eysenck & Eysenck, 1960; Eysenck *et al.* 1962). Extraverts, in whom, according to the individual postulate of Eysenck's theory, cortical inhibition develops rapidly, have been found to have shorter after-effects than introverts, in whom cortical inhibition develops slowly (Eysenck *et al.* 1957; Holland, 1960; Eysenck, 1960; Claridge, 1960; Lynn, 1960).

On the basis of these findings and the theory of homoeostatic excitation, it was predicted that under conditions of increased cortical inhibition an increase of duration of stimulation would produce a shortening of the spiral after-effect duration and that this reduction in the after-effect would be greater for the expansion after-effect than for the contraction after-effect. The argument underlying this prediction is that the level of inhibition accumulated with increase of stimulation would be greater for the expansion after-effect since the inhibition would be retained within the contour of the spiral disk.

Similarly, it was predicted that under conditions of increased cortical excitation the increase in duration of the after-effect produced by increase in the duration of stimulation would be greater for the expansion after-effect than for the contraction after-effect.

More specifically the predictions tested in Expt. II were: (1) The expansion after-effect of extraverts produced by 120 sec. stimulation will be significantly shorter than that produced by 30 sec. stimulation. (2) The contraction after-effect of extraverts will show no change with increase of stimulation. (3) The expansion after-effect of introverts produced by 120 sec. stimulation will be significantly longer than that produced by 30 sec. stimulation. (4) The contraction after-effect of introverts will show no change with increase in duration of stimulation.

Method

The subjects used were 12 boys and 12 girls whose ages ranged from 14 to 17. They were chosen from a group of 250 children on the basis of their scores on the Junior M.P.I. (Furneaux & Gibson, 1961) such that their scores fell at least 1 S.D. above or below the mean for the total group.

Each of the groups of 12 extraverts and 12 introverts was then split on a random basis into four groups of three subjects each. Each of these four groups received one of four different combinations of direction of rotation and duration of stimulation.

The subject was given a 15 sec. practice trial on the spiral and after a rest period of 15 min. was given one of the four following stimulating conditions: (1) 30 sec. clockwise rotation; (2) 120 sec. clockwise rotation; (3) 30 sec. counterclockwise rotation; (4) 120 sec. counterclockwise rotation.

Results

The mean duration of the after-effect for each of the eight groups is shown in Table 2.

Table 2. *Mean duration of the spiral after-effect in sec. (Expt. II)*

		Duration of stimulation	
Group	Type of after-effect	30 sec.	120 sec.
Extraverts	Expansion	15·40	27·13
	Contraction	11·40	38·27
Introverts	Expansion	8·67	20·80
	Contraction	5·07	17·53

An analysis of variance of the data (Table 3) indicated: (1) that the main effects of extraversion and duration of stimulation were significant; (2) that the second-order interaction between direction of rotation, duration of stimulation and extraversion was significant.

Contrary to the prediction from Eysenck's theory and out of line with previous findings, the after-effect of the extraverts was significantly longer than that of introverts. No satisfactory explanation has been found for this result.

For both extraverts and introverts the expansion and contraction after-effects showed a significant increase from 30 to 120 sec. stimulation. The significant second-order interaction would appear to be due to the greater increase in the case of the contraction after-effect of the extraverts.

Only the significant increase of the introverts' expansion after-effect with increase in stimulation time is in line with prediction (Prediction 3). The other findings are quite contrary to prediction.

Table 3. *Analysis of variance of the data obtained on extraversion–introversion, direction of rotation and duration of stimulation (Expt. II)*

Source	Sum of squares	D.F.	M.S.V.	P
Due to differences in direction	0·02	1	0·02	—
Due to differences in duration of stimulation	1497·84	1	1497·84	< 0·001
Due to differences in extraversion–introversion	604·00	1	604·00	< 0·001
Due to direction : duration interaction	89·71	1	89·71	—
Due to direction : extraversion interaction	23·51	1	23·51	—
Due to duration : extraversion interaction	73·50	1	73·50	—
Due to interaction between direction, duration and extraversion	132·14	1	132·14	< 0·05
Residual error	468·80	16	29·30	
Total	2889·52	23		

Experiment III

Despite the lack of confirmation of the predictions, it was decided to make a further test of the predicted effects of increasing inhibition or excitation. In the third experiment, level of inhibition was increased by the use of depressant drugs (sodium seconal and meprobamate) and level of excitation was increased by the use of a stimulant drug (dextro-amphetamine).

The assumption that these drugs would have the stated effects on level of inhibition and excitation was based on Eysenck's drug postulate which states that 'depressant drugs increase cortical inhibition and decrease cortical excitation and thereby produce extraverted behaviour patterns whereas stimulant drugs decrease cortical inhibition, increase cortical excitation, and thereby produce introverted behaviour patterns' (Eysenck, 1961).

On the basis of the argument presented for Expt. II relating level of inhibition and excitation to direction of rotation and duration of stimulation, it was predicted: (1) The expansion after-effect of subjects who had received sodium seconal or meprobamate would be significantly shorter after 120 sec. stimulation than after 30 sec. stimulation. (2) The contraction after-effects of the subjects who had received sodium seconal or meprobamate would show no change with increase in duration of stimulation from 30 to 120 sec. (3) The expansion after-effects of subjects who had received dextro-amphetamine would be significantly longer after 120 sec. stimulation than after 30 sec. stimulation. (4) The contraction after-effect of subjects who had received dextro-amphetamine would show no change with an increase of stimulation from 30 to 120 sec.

Method

The subjects used were 74 females and 26 males whose ages ranged from 17 to 43. They were chosen from a group of 331 subjects (244 females and 87 males) on the basis of their scores on the Maudsley Personality Inventory (Eysenck, 1959). To rule out the effects of the extraversion variable, only those subjects were used whose extraversion scores were less than 1 S.D. from the mean for the total group.

The selected group of 100 subjects was divided randomly into 20 subgroups of five subjects each. Each of these subgroups received one of 20 different combinations of five drug treatments, two directions of rotation and two durations of stimulation.

Each subject was given a 15 sec. practice trial on the spiral. One and a half hours after administration of the drug treatment, one of the four following stimulating conditions with the spiral were presented: (1) 30 sec. clockwise rotation; (2) 30 sec. counter-clockwise rotation; (3) 120 sec. clockwise rotation; (4) 120 sec. counter-clockwise rotation.

The subject received one of the following drug treatments: (1) no drug (the subjects in this group were tested 1½ hr. after the practice session); (2) placebo; (3) 10 mg. dextro-amphetamine; (4) 800 mg. meprobamate; (5) 195 mg. sodium seconal.

Table 4. *Mean duration of the spiral after-effect in sec.* (*Expt. III*)

	Expansion (sec.)		Contraction (sec.)	
	30	120	30	120
No drug	25·38	50·02	18·88	19·19
Placebo	28·43	21·16	17·51	21·46
Meprobamate	20·99	38·23	14·31	27·00
Sodium seconal	28·85	18·19	11·94	39·65
Dextro-amphetamine	31·93	23·74	17·67	20·66

Table 5. *Analysis of variance of the data obtained on direction of rotation, duration of stimulation and drug administered* (*Expt. III*)

Source	D.F.	Sum of squares	M.S.V.	P
Due to direction of rotation	1	1,548·50	1,548·50	< 0·01
Due to drug	4	431·66	107·91	—
Due to duration of stimulation	1	1,003·68	1,003·68	< 0·05
Direction : duration interaction	1	255·08	255·08	—
Direction : drug interaction	4	1,138·25	284·56	—
Drug : duration interaction	4	1,305·71	326·43	—
Due to interaction between direction, drug and duration	4	2,666·07	666·52	< 0·025
Residual error	80	15,293·50	191·17	
Total	99	23,542·45		

Results

The mean duration of the after-effect for the 20 subgroups is shown in Table 4.

An analysis of variance of the data (Table 5) indicated that: (1) The main effect of direction of rotation was significant. (2) The main effect of duration of stimulation was significant. (3) The second-order interaction between direction of rotation, duration of stimulation and drug treatment was significant.

In line with previous findings, the expansion after-effect was significantly longer than the contraction after-effect.

In line with prediction from the theory of homoeostatic excitation–inhibition and the findings of Expt. I, there was a significant increase in the expansion after-effect and no change in the contraction after-effect under no drug conditions.

In line with prediction (1) there was a decrease in the expansion after-effect with increase of stimulation for the sodium seconal group. But this difference is not statistically significant when tested by a t test based on the analysis of variance.

Contrary to prediction (2) there was a significant increase in the contraction after-effect with increase of stimulation for the sodium seconal group. It would appear, however, to be in line with the significant increase in the contraction after-effect with an increase of stimulation time found with the extraverts in Expt. II.

Contrary to prediction (3) there was a drop in duration of the expansion after-effect with increase of stimulation for the dextro-amphetamine subjects though this drop is not statistically significant.

In line with prediction (4), the contraction after-effect of the dextro-amphetamine group showed little change with increase in duration of stimulation.

It will be noted that, for the placebo group, the expansion after-effect did not increase with increased stimulation. No explanation for this result can be given. The effects of meprobamate were not significant.

EXPERIMENT IV

On the basis of his theory of cortical excitation–inhibition, Eysenck has predicted that massed practice should lead to a reduction in the length of the spiral after-effect. He argues that the neural pathways mediating the after-effect are prone to inhibition or satiation and that massed practice should lead to an accumulation of inhibition. He also has predicted that a period of rest should lead to a dissipation of inhibition and thus to reminiscence, i.e. a lengthening of the after-effect to near its original value (Eysenck & Eysenck, 1960).

The predictions have not received unequivocal confirmation. Only the studies by Eysenck & Eysenck (1960), and by Eysenck *et al.* (1962), present data in line with the predicted decrease in the after-effect with massed practice. The studies by Eysenck *et al.* (1957), and Pickersgill & Jeeves (1958), though showing a drop in duration of after-effect with repeated stimulation, are concerned with repeated trials rather than massed trials. Costello (1961) found no significant effects of massed practice in one study, and in a second study (Costello, 1963) found a significant decrease followed by significant increase in duration of the after-effect.

With regard to reminiscence, there is even less substantiation of the predicted effect since there is no evidence of a significant reminiscence effect in any of the above studies and in one study (Costello, 1963) there was a significant decrease in the after-effect after a rest period.

An attempt has been made to explain the significant decrease followed by a significant increase in duration of the after-effect in Costello's study, by means of the theory of homoeostatic excitation–inhibition. It was proposed that the inhibition produced by massed practice builds up to a critical level beyond which further massed practice will result in an increase in excitation up to a second critical level beyond which inhibition will once more appear and so on.

Though in the 1961 study Costello did not find a significant effect of massed practice during the first six massed trials, during the second set of six massed trials presented after the rest period of 10 min. he found that the contraction after-effect showed a significant drop in duration from trial 1 to trial 3 followed by a significant increase from trial 3 to trial 6. After the rest period, the expansion after-effect alternately increased and decreased by a statistically insignificant amount.

The purpose of this experiment was to investigate further the effects of massed practice on the spiral after-effect and the interaction of this effect with the effect of difference in direction of rotation of the spiral and the effect of the administration of stimulant and depressant drugs.

In the introduction to Expt. II we gave our reasons for believing that inhibition accumulates more readily in the case of clockwise rotation of the spiral. It was predicted, therefore, that massed practice would result in alternating decrease and increase of the spiral after-effect and this would be more marked in the case of the expansion after-effect.

It was further predicted that this effect of massed practice would be more marked after administration of depressant (sodium seconal, meprobamate) or stimulant drugs (dextro-amphetamine) since these drugs would produce an increase in the level of cortical inhibition or cortical excitation.

Method

The subjects used were the 74 females and 26 males used previously for Expt. III. The selected group of 100 subjects was divided randomly into 10 subgroups of 10 subjects each. Each of these subgroups received one of 10 different combinations of five drug treatments and two directions of rotation. Each subject received the same drug treatment he had received in Expt. III.

The subject was given a 15 sec. practice trial on the spiral. At the end of the rotation period he was to watch the after-effect and to press a switch on his right when the after-effect ceased. One and a half hours after administration of the drug treatment, he was given 10 massed trials followed by a 15 min. rest period and three further massed trials. Half the subjects were presented with a clockwise rotation (expansion after-effect) and half with a counter-clockwise rotation (contraction after-effect).

Results

The results for the no drug, sodium seconal and dextro-amphetamine groups are presented in Fig. 1.

The results for the meprobamate and placebo groups have been omitted from the graph since they were similar to the results for the no drug group. The results for all the groups were, however, included in the analysis of variance which followed the method proposed by Eysenck *et al.* (1962).

With ten groups of ten subjects each subjected to experimentation with a particular direction and drug and ten trials made before the rest pause and three after, the total number of observations is

$$10 \times 10 \times 13 = 1300.$$

The total observed variation about the general mean is 110289·19 with 1299 D.F. It is separable into two parts, variation among the subjects and variation within the subjects. The latter part consists of each subject's variation about his own mean for the 13 trials so that it has 100 × 12 = 1200 D.F. It can be separated again into two parts: variation among trials with 12 D.F. and a residual due to interaction between person and trial.

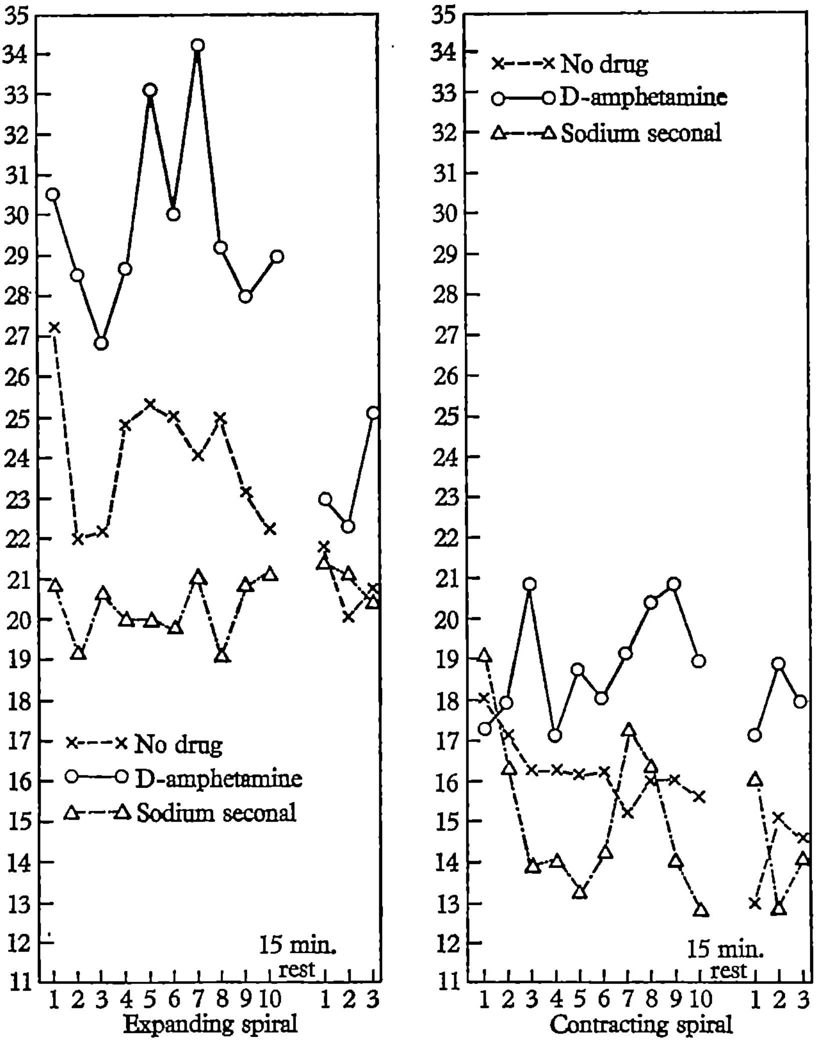


Fig. 1. Spiral after-effects as a function of drugs administered, direction of rotation and massing of practice. (Length of after-effects in seconds is shown on the ordinate; trials and rest periods are shown on the abscissa.)

The first part of the analysis, concerning variation among the subjects, is shown in Table 6.

Differences due to differences in direction of rotation are highly significant. The other effects are not significant.

The second part of the analysis, concerning variation among trials, is also shown in Table 6.

It can be seen that the effect of the pause is statistically significant.

Finally, the residual variation within the subjects can be analysed as shown in Table 6.

It can be seen that only the interaction between direction of rotation and trials reaches a statistical level of significance.

Table 6

Source	Sum of squares	D.F.	M.S.V.
Variation among the subjects			
Due to differences in direction	12,074·22	1	12,074·22**
Due to differences in drug	5,269·04	4	1,317·26
Due to direction : drug interaction	1,729·76	4	432·44
Residual error, i.e. variation between subjects under the same experimental conditions	60,211·11	90	669·01
Total observed	79,284·13	99	
Variation among the trials			
Due to the pause	1,476·87	1	1,476·87**
Residual variation between trials not separated by the pause	383·82	11	34·893
Total observed	1,860·69	12	155·06
Residual variation			
Due to drug: trial interaction	1,344·35	48	28·01
Due to direction: trial interaction	590·38	12	49·20*
Due to interaction of direction : drug and trials	1,170·67	48	24·39
Not associated with any experimentally controlled effects	26,038·97	1080	24·11
Total observed	29,144·37	1188	24·53

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$.

Although combining all groups there is an over-all drop in duration of the after-effect from trial 1 (mean 21·35 sec.) to trial 10 (mean 19·92 sec.), it can be seen from the graph that there is an alternating decrease and increase in the duration. This is in line with the general prediction from the theory of homoeostatic excitation–inhibition. This effect of massing trials would seem to be more in evidence for the expansion than for the contraction after-effect and this is substantiated by the significant interaction of direction of rotation and trials effect. This also is in line with prediction.

Though examination of the graph suggests that there is an interaction between direction of rotation, trials and drug, this interaction does not reach statistical significance and the prediction regarding the effects of the drugs is not confirmed. The effect of the pause was significant though there was a drop in duration of the after-effect following the pause contrary to Eysenck's predicted reminiscence effect. The over-all mean for trial 10 was, as noted above, 19·92 sec. and for the first post-pause trial 18·03 sec.

DISCUSSION

The predictions from the theory of homoeostatic excitation–inhibition that there would be a significant increase in the expansion after-effect with an increase in stimulation time and that the contraction after-effect would show no change with increase of stimulation time were confirmed in Expt. I.

The general prediction of a second-order interaction between extraversion, direction of rotation and duration of stimulation was confirmed by Expt. II. But the more specific predictions concerning this interaction were not confirmed—the results being quite contrary to prediction.

Similar results were obtained with the predicted second-order interaction between drugs, direction of rotation and duration of stimulation in Expt. III.

In Expt. IV the predicted alternating decrease and increase of the after-effect with massed trials did occur and was more in evidence for the expansion after-effect. The prediction that this effect of massed practice would be increased after administration of depressant or stimulant drugs was not confirmed.

One is tempted to conclude that the theory has received little confirmation and must be reformulated. But is this really the case? The main proposition of the theory is that accumulation of excitation may result in homoeostatic inhibition and accumulation of inhibition may result in homoeostatic excitation. Since the theory postulates that 'paradoxical' reversal of the state of the central nervous system may occur by its very nature, the occurrence of 'paradoxical' behavioural observations is predicted.

The theory then has, in one sense, been supported by the findings. But by its very confirmation it follows that the theory is useless. It is useless so far as specific predictions are concerned.

It may be argued that the theory simply needs reformulating to account for the findings and to generate specific predictions. It appears, however, that reformulation would have to be to a degree such that the whole nature of the theory would be changed. The findings suggest that molar neuropsychological theories may be of little value and that theories concerning psychological and neural events would have to be formulated at a molecular level in relation to specific functions.

One is often presented with the argument of the molecular theorists that it is impossible to predict performance on, for instance, the spiral after-effect because there are too many variables involved. They argue that one would have to study intensively the parameters of the spiral after-effect phenomenon before attempting to make predictions concerning the function.

The findings reported here have suggested that homoeostatic or seemingly paradoxical events occur readily in the nervous system and gives the molecular theorists a further argument for their case.

This work was supported by a grant from Wallace Laboratories, New Jersey. Thanks are due to Mrs Gloria Glab, who conducted the experiments, to the medical staff of the Munroe Wing for their supervision of drug administration, and to Dr Patrick Slater for his statistical advice.

REFERENCES

CLARIDGE, G. (1960). The excitation-inhibition balance in neurotics. In EYSENCK, H. J. (ed.), *Experiments in Personality*. London: Routledge & Kegan Paul.

COSTELLO, C. G. (1960). Further observations on the spiral after-effect. *Percept. Mot. Skills*, **11**, 324.

COSTELLO, C. G. (1961). Effects of massed practice on the spiral after-effect and the homeostatic nature of excitation-inhibition. *Percept. Mot. Skills*, **12**, 11–14.

COSTELLO, C. G. (1963). The effects of meprobamate on the spiral after-effect. In EYSENCK, H. J. (ed.), *Experiments with Drugs*. London: Pergamon.

DEUTSCH, J. A. (1956). The statistical theory of figural after-effects and acuity. *Brit. J. Psychol.* **47**, 208–15.

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory.* London: University of London Press.

EYSENCK, H. J. (1960). Objective psychological tests and the assessment of drug effects. *Int. Rev. Neurobiol.* II.

EYSENCK, H. J. (ed.) (1961). *Handbook of Abnormal Psychology.* New York: Basic Books Inc.

EYSENCK, H. J. & EYSENCK, S. B. G. (1960). Reminiscence on the spiral after-effect as a function of length of rest and number of pre-rest trials. *Percept. Mot. Skills*, **10**, 93–4.

EYSENCK, H. J., HOLLAND, H. C. & TROUTON, D. S. (1957). Drugs and personality. III. The effects of stimulant and depressant drugs on visual after-effects. *J. Ment. Sci.* **103**, 650–5.

EYSENCK, H. J., WILLETT, R. A. & SLATER, P. (1962). Drive, direction of rotation, and massing of practice as determinants of the duration of the after-effects from the rotating spiral. *Amer. J. Psychol.* **75**, 127–33.

FURNEAUX, W. D. & GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.* **31**, 204–7.

GALLESE, A. S. (1956). Spiral after-effect as a test of organic brain damage. *J. clin. Psychol.* **12**, 254–8.

HOLLAND, H. C. (1960). The effects of depressant drugs on perceptual processes. In EYSENCK, H. J. (ed.), *Experiments in Personality.* London: Routledge and Kegan Paul.

HOLLAND, H. C. & BEECH, H. R. (1958). The spiral after-effect as a test of brain damage. *J. Ment. Sci.* **104**, 466–71.

LYNN, R. (1960). Extraversion, reminiscence and satiation effects. *Brit. J. Psychol.* **51**, 319–24.

PHILBRICK, E. B. (1959). The validity of the spiral after-effect as a clinical tool for diagnosis of organic brain pathology. *J. Consult. Psychol.* **23**, 39–43.

PICKERSGILL, M. J. & JEEVES, M. A. (1958). After-effect of movement produced by a rotating spiral. *Nature, Lond.*, **182**, 1820.

SPIVACK, G. & LEVINE, M. (1957). The spiral after-effect and reversible figures as measures of brain damage and memory. *J. Personality*, **25**, 767–78.

(*Manuscript received* 8 *April* 1963)

Brit. J. Psychol. (1964), 55, 2, pp. 201-206
Printed in Great Britain

EFFECTS OF SENSORY INPUT AND SENSORY DOMINANCE ON SEVERELY DISTURBED, AUTISTIC CHILDREN AND ON SUBNORMAL CONTROLS

By BEATE HERMELIN AND N. O'CONNOR

Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry, London, S.E. 5

Sensory dominance in children differs with age, development and experience and may be a diagnostic differential in the case of autistic children. This hypothesis was explored in three experiments with autistic children, one concerned with sensory dominance, one with sense discrimination learning and one with the dynamics of this learning process. Imbecile children of the same performance I.Q. were used as controls.

The experiments showed that psychotic and control children responded most readily to light stimuli, but that this was followed by response to touch stimuli in the autistic children, and by response to sound stimuli in the controls. Differential rewarding resulted in dominance changes to the same extent in each group. However, autistic children did not learn to suppress a previously dominant stimulus orientation habit as quickly as the controls.

I. INTRODUCTION

There is some evidence that the concept of a hierarchical structure of sensory systems is connected with the developmental process in children. Amongst others Renshaw (1930), Zaporozhets (1961) and Birch (1962) used this concept. According to this theory interoceptive and visceral sensations are dominant in the young child, and this dominance is gradually superseded, first by tactile and kinaesthetic and then by the telereceptors, i.e. the auditory and visual sensory systems. A parallel process is the gradual establishment of the equivalence and integration of different sensory information, so that stimuli to one sense can be readily recognized and interpreted in another. Such concepts are of interest for investigating responses to environmental stimulation in psychotic or autistic children. Goldfarb (1956) has suggested that schizophrenic children make contact with their environment through the proximal rather than the distance receptor channels. The relative dominance of different receptors may be important for the control of the behaviour of such children in an experimental environment.

Three experiments were therefore planned to study the effects of sensory input and dominance in severely disturbed autistic children. The first experiment studied sensory dominance, the second sensory discrimination learning and the third analysed the dynamics of this learning process.

II. EXPERIMENT I

Subjects

Severely subnormal disturbed children were selected by a psychiatrist on the basis of the presence of manneristic or autistic behaviour. Controls were matched for age, sex and as far as possible for I.Q. Each subject was then checked against a list of behaviour items, selected by Creak *et al.* (1961), which describes behaviour frequently found in autistic children. No child was selected for the autistic group unless he had at least three of these symptoms and either one of two items regarded as of major diagnostic importance. No control showed any symptom. The selection of subjects was made on the basis of behavioural criteria only. Children with evidence of C.N.S. damage were included in both the experimental and control groups along with other children without such signs. No difference between the groups was found concerning frequency of such clinical signs.

The mean age of the psychotic and control groups was 12 years, ranging from 8·6 to 16 years. The mean I.Q. for the psychotic children was 40 points, ranging from 28 to 55 and that for the controls 43 points with a range from 30 to 56.

Method

Ten autistic and ten undisturbed, severely subnormal children acted as subjects. The aim was to isolate those sense avenues to which the subjects would be pre-eminently responsive in the organization of their response behaviour.

Subjects were tested individually whilst sitting at a small table opposite the experimenter with the apparatus on the table between them. Simultaneous bimodal stimulation was presented, and visual, auditory and tactile stimuli were used. Lights were placed on either end of a 30 in. by 15 in. hardboard display panel, and buzzers were mounted behind the screen, out of sight of the subject. Loops on a piece of string, secured around the ankles on each leg of the subjects and pulled by the experimenter, provided tactile stimulation. There were six possible stimulus combinations:

Light (right)/Sound (left	Sound (right)/Light (left)
Light (right)/Touch (left)	Touch (right)/Light (left)
Touch (right)/Sound (left)	Sound (right)/Touch (left)

Thus at each trial two stimuli of different modalities were presented simultaneously, one on the left and one on the right side. Order of presentation of stimulus combinations was randomized over 60 trials by the use of a Latin Square. Thus each stimulus combination occurred ten times. Apart from a subjective judgement by the experimenters no attempt was made to equate intensities of the stimuli over different modalities.

Prior to the experimental trials each subject was presented with all stimuli singly and in succession. He was instructed to point to the side at which the stimulus was presented, and was then rewarded with a sweet. When, as was the case with some of the autistic children, the subject refused to point, turning of the head or eyes towards the stimulus source was accepted as response.

Following the training session, and without further instruction, two stimuli in different modalities were presented simultaneously, one on the right and the other on the left side. Thus if, for instance, the subject received a tactile stimulus on the left together with a visual one from the right, he could point to either side, and each response, regardless of to which stimulus it occurred, was rewarded. Each response was recorded.

Results

The data were investigated for the following trends. First, it may have been that position habits had developed, for instance the subject always pointing to the right, or alternating between left and right. As each of the stimuli occurred equally frequently on either side, an equal distribution of scores would thus have resulted over modalities. No score of any child approached such a pattern. Similarly, if position were the determining factor in choice, then the scores in, for instance, Sound (right)/Light (left), and Light (left)/Sound (right) conditions should differ. Since t tests revealed no such differences, the six conditions were combined to give three combinations of Sound/Light, Light/Touch and Sound/Touch for further analysis. The total numbers of responses made to each stimulus under each condition are shown in Table 1.

The table shows clearly that both groups responded predominantly to light whenever a visual stimulus was a component in the combination of signals. It also shows that, when touch and light were combined, the autistic children though still obtaining a clearly dominant light score, responded somewhat more frequently to the touch stimulus than did the controls. However, the most marked difference between the groups is evident in the conditions in which a sound and a touch stimulus were the alternatives. In this instance, the controls showed a clearly dominant sound score, whereas the responses of the psychotics were made more often to the touch stimulus. If the frequency distributions of light versus sound responses are compared, the response pattern of the groups does not differ significantly. Subjects from both groups respond more frequently to light than sound. However, there is a highly significant difference between the groups in the distribution of scores in the Sound–Touch stimulus combination ($t = 9{\cdot}67$, $P < 0{\cdot}001$). As stated, this reflects the touch dominance in the psychotics in contrast to the sound dominance in the controls. The relative predominance of tactile input for the psychotics is also evident in the pattern of responses to the Light–Touch stimulus combination. The difference in the distribution of frequencies between the groups is significant ($t = 2{\cdot}88$, $P < 0{\cdot}02$).

Table 1. *Number of responses to bimodal stimulus presentation*

Stimulus conditions	Number of responses (optimal 200)	
	Psychotics	Controls
Sound	26	18
Light	174	182
Touch	48	2
Light	152	198
Sound	53	168
Touch	147	32

III. EXPERIMENT II

The same 20 children who acted as subjects in the first experiment took part in a series of discrimination learning trials. But while every response, regardless to which stimulus it had been made, was rewarded in Expt. I, only the selection of either the sound or the touch was rewarded in this series. Two questions were asked. First, would there be a difference between the groups in the ability to reject a previously preferred stimulus and choose instead one formerly less favoured? Secondly, would there be a differential effect of sound and touch on the response pattern of the subjects, the psychotic finding it easier to learn to respond to tactile, and the controls to auditory stimulation?

Method

The same apparatus was used as that in the previous experiment. Light was presented simultaneously with either sound or touch for 40 trials, 20 for each condition. Presentation for stimulus combinations and left or right positions was varied in a balanced random order.

The child was told that he would be given a counter every time he pointed to the correct one of two stimuli, and that five counters would be exchanged for a sweet. The selection of either sound or touch was rewarded while that of light was not.

Results

The score for each subject was the number of correct responses to the sound or touch stimulus, while ignoring the light on the opposite side of the display panel. These scores are given in Table 2.

No significant difference was found for either group in the frequency with which they responded to sound or touch. The total number of correct responses also did not differ significantly between the groups, nor did the number of children from either group who gave more than 20 correct responses out of a possible 40. An analysis of variance was carried out, which compared the frequency of responses to light in the first and second experiments. As previously stated, both groups gave significantly more responses to light than to either sound or touch in the previous experiment, when the responses were rewarded regardless to which signal modality they were made. As responses to light remained unrewarded in the second experiment, a decrease of such responses in comparison with Expt. I can be regarded as a function of the changed re-inforcement schedule. A summary of the results of the analysis is given in Table 3.

As can be seen, there was a highly significant difference between experiments, confirming the effect of re-inforcement on response to different stimulus input. This was the case for both groups.

Table 2. *Number of correct responses to sound and touch stimuli*

	Psychotics	Controls
Touch	92	106
Sound	103	121

Table 3. *Analysis of variance, dominance versus discrimination scores*

Source	D.F.	Sum of squares	Mean square	*F*	*P*
Between groups	1	12·1	12·1	—	—
Between treatments	1	2822·4	2822·4	28·6	0·001
Groups/treatments	1	184·9	184·9	—	—
Between people within groups	18	1710·25	95·0	—	—
Residual	18	1776·7	98·7	—	—
Total	39	—	—	—	—

IV. EXPERIMENT III

The difference in response pattern following a change in the re-inforcement schedule, which had been found in psychotic children as well as in controls, was due to decreasing a predominance of responses to light stimuli found in the first experiment to a level at which responses to either of the three sensory stimuli were equally frequent in Expt. II. This might have resulted from an increased dominance of sound or touch as opposed to light or from an increased capacity to ignore light in favour of any other stimulus whatever. In order to analyse further the effect of different sensory input on the response pattern of the two groups, the task was broken up into its components, i.e. responses to touch and sound, and inhibition of such a response to light.

Method

Two children from each group of those previously used as subjects were not available for retesting. The other 18 were used again, as well as six more, who had not previously taken part. There were thus 12 subjects diagnosed as psychotic and 12 subnormal controls. Each subject was again given 40 stimulus presentations. However, where previously two stimuli of different sense modality had been presented simultaneously, this time light, sound and touch were presented singly. Twenty trials consisted of presenting sound or touch on either the left or the right side, and responses to the side on which the stimulus occurred were rewarded. There were also 20 trials in which the light alone was presented on either the right or the left, and in those trials the response to the opposite side, i.e. the one on which the light was *not* was rewarded. Tasks I and II were given first and second to alternate subjects. Presentation of the stimulus on either the right or the left side, and sequence of sound and touch stimuli were varied in a balanced random order.

Results

The results for both tasks and both groups are given in Table 4.

Table 4. *Number of orientation responses to sound and touch and inhibition of responses to light*

	Psychotics	Controls
Orientation	171	185
Inhibition	75	182

The significant interaction in the analysis of variance, summarized in Table 5, shows that while there is no difference between the two tasks in the controls, there is such a difference in the psychotic group. While the psychotic children do not differ from the controls in their ability to respond to the side on which sound or touch are presented, they do differ from them in the frequency of responses to that side of the display opposite to that on which the light is presented. Because of the significant interaction term in the analysis, main terms are not discussed.

Table 5. *Analysis of variance of orientation and inhibition responses*

Source	D.F.	Sum of squares	Mean square	*F*	*P*
Between groups	1	305·02	305·02	11·67	—
Between treatments	1	204·18	204·18	9·76	—
Groups/treatments	1	180·19	180·19	8·62	0·01
Between people within groups	22	574·96	26·13	—	—
Residual	22	460·13	20·91	—	—
Total	47	—	—	—	—

V. DISCUSSION

The results obtained in these three experiments may be summarized by saying that when presented with bimodal combinations of light, sound and touch stimuli, psychotic children as well as their subnormal controls responded most frequently to light. When a stimulus combination of sound and touch was presented, the psychotics tended to respond to touch and the controls to the sound stimulus. However, when in the second experiment responses to sound or touch only were rewarded, while those to light were not, the number of responses to light decreased in both groups equally.

Thus while hierarchical dominance of the visual systems had developed in both groups, the pattern of response behaviour to sound and touch indicated differential effects of sensory input for the two groups. This lack of dominance of auditory over tactile stimulation in the psychotic children may be related to their very impaired speech development.

When touch and light were presented simultaneously, one on the right and the other on the left, and then response to touch was rewarded while response to light was not, the frequency of response to light and touch were equal. The same is true of the simultaneous light-sound combination and these results were found in both the groups. This finding suggests that psychotic children are affected by changes in re-inforcement schedules just as controls are. It supports the results of Ferster & DeMyer (1962) who were able to control the behaviour of autistic children through instrumental conditioning techniques. These authors concluded that the results of such techniques did not suggest any basic deficit in the rate at which these children acquired new patterns of behaviour, when compared with normals. When, as in the present study, the effects of re-inforcement on behaviour are compared between psychotic children and subnormal children without behaviour disturbance, both groups respond to the same extent. In the present study a spontaneous preference, i.e. the response to light, had to be extinguished in favour of other re-inforced stimuli. Our results show that it is this extinction of a dominant response pattern, rather than the establishing of a new one, which is impaired in psychotic children. While both processes were controlled to the same extent through the re-inforcement schedule in the controls, there was a differential effect in the experimental group. Lack of inhibition of dominant behaviour patterns, rather than the ability to acquire new ones, may distinguish psychotic children from their subnormal controls.

Our thanks are due to Dr J. M. Crawford, Dr D. W. Smith and Mrs K. Mair of Botleys Park Hospital and to Dr Gibson and Dr P. Sylvester of St Lawrence's Hospital, and to the staff of these hospitals for their kind co-operation.

REFERENCES

BIRCH, H. G. (1962). Dyslexia and the maturation of visual function. In MOREY, J. (ed.), *Reading Disability*. Baltimore: Johns Hopkins Press.

CREAK, MILDRED, *et al.* (1961). Schizophrenic syndrome in children. *Lancet*, **2**, 818.

FERSTER, C. B. & DEMYER, M. K. (1962). A method for the experimental analysis of the behaviour of autistic children. *Amer. J. Orthopsychiat.* **32**, 89–98.

GOLDFARB, W. (1956). Receptor preference in schizophrenic children. *Arch. Neurol. Psychiat.* **76**, 643–52.

RENSHAW, S. (1930). The errors of cutaneous localization and the effect of practice on the localizing movement in children and adults. *J. Gen. Psychol.* **38**, 223–38.

ZAPOROZHETS, A. V. (1961). The origin and development of the conscious control of movements in man. In O'CONNOR, N. (ed.), *Recent Soviet Psychology*. London: Pergamon.

(*Manuscript received* 27 *April* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 207–218
Printed in Great Britain

PERFORMANCE AND LEVEL OF ACTIVATION IN SCHIZOPHRENICS AND NORMALS

BY P. H. VENABLES
Social Psychiatry Research Unit, Medical Research Council,
Institute of Psychiatry, Maudsley Hospital, London, S.E. 5

Four experiments are presented which provide evidence for a relationship between performance and indices of level of activation. This relationship takes the form of a number of equally spaced U's or V's rather than the single U which is expected from the literature. Mechanisms which might be responsible for producing this pattern of results are discussed.

I. INTRODUCTION

While an inverted U relationship between activation or arousal level and performance has been proposed (Hebb, 1955; Malmo, 1959), not many attempts have been made to provide direct evidence of its validity. Prior to the introduction of arousal or activation concepts into psychological literature (Lindsley, 1951) workers such as Duffy (1951) and Freeman (1948) had suggested an inverted U-shaped relationship between indices of 'behavioural intensity' and performance. Experiments by, for example, Freeman (1940) and Courts (1942) provided evidence to support this form of relationship. More recently Stennett (1957) and Furth & Terry (1961), for example, related complex performance and learning to physiological indices of arousal or activation in an inverted U-shaped fashion. Attempts, however, to repeat Freeman's original (1940) experiment on simple RT and skin conductance level have not met with success (Schlosberg & Kling, 1959). These workers suggest that this failure is due to the lack of use of multiple indices of activation level. On the basis of the results to be presented here it is proposed that the failure to find a simple U-shaped relationship could occur for three reasons. First, that the relationship in fact takes the form of a number of U's and that the performance/activation relationship of any one subject does not necessarily remain on the same 'U' on successive periods of testing. Secondly, performance and activation level are not related at certain performance levels. Thirdly, that measures of autonomic activity level may not necessarily be relevant where it is necessary to know the state of activity at a cortical level.

The momentary state of activation at the cortex measured by the state of the alpha rhythm has been related to simple RT by Walsh (1952), O'Hare (1954), Lansing (1957), Lansing, Schwartz & Lindsley (1959), Fedio, Mirsky, Smith & Parry (1961), Callaway (1962) and Surwillo (1963). The absence of a satisfactory index of the relatively more permanent *level* of activation of the cortex has made it difficult to provide an adequate test of the validity of the inverted U hypothesis when relating performance to cortical activation level. It has been proposed (Venables, 1963*a*) that the threshold at which the subject reports fusion of a pair of temporally close light flashes may be used as an index of cortical activity level. A first experiment, relating on a single occasion this index to simple reaction time using schizophrenics as subjects, gave strong evidence of a multiple V relationship between the two measures. In view of the rather disturbingly close fit of the data to what was virtually an arbitrary set of parameters, a second experiment was performed with the following aims. First, to see whether the data fitted the parameters of the multiple V pattern set up by the first experiment. Secondly, to examine the data from the same subject on three separate occasions to see whether the subject kept to the 'same V' as required by Freeman's (1940) experiment or whether the data from one subject required a multiple V relationship to encompass it. Thirdly, as the first experiment had used only schizophrenics as subjects, to see whether normal subjects showed results fitting the same set of parameters. Fourthly, in view of the close relationship (Venables, 1963*a*) between threshold of fusion of two flashes and skin potential level as indices of activation level in schizophrenics, to see whether the data would fit the multiple V relationship previously found directly transcribed into skin potential terms. This final aim was included because it was thought that the multiple V relationship could possibly result from an artifact in the measurement of two flash thresholds. It was unlikely that the same artifact could be present in a measure so different as skin potential.

To test further the generality of the multiple V relationship, experiments were undertaken in which performance on card sorting and a free association task was related to cortical activation.

II. METHOD

Apparatus

(*A*) *Measurement of threshold of fusion of paired light flashes*

Approximately square wave flashes of light 5 msec. in duration were produced by applying pulses to the grid of a Ferranti CL. 66 cathode-ray tube. This tube has a phosphor coating which produces an essentially white source of light and has a decay time of 5 μsec. Associated timing equipment made possible the presentation of two flashes of light the interval between which was continuously variable and could be set to the nearest millisecond in the working range of the experiment, between 30 and 140 msec. The intensity of the flash was attenuated to 50 ft. lambert by a neutral density filter. An edge-lit Perspex block in front of the tube face provided a fixation patch of 0·025 ft. lambert.

A mask with a circular hole in front of the tube face gave a visual angle of 1·8°. The experimental room was completely darkened and the timing apparatus was operated from an adjacent room. A simple press button signalling system enabled the subject to report the number of flashes he saw.

The subject was seated with his eyes approximately 4 ft. from the tube face, he was asked to rest his head between two pads on the chair back which kept it in line with the optical axis of the apparatus. He was told that he would see either one or two flashes of light and that all he had to do was to report the number by giving one or two presses on a button conveniently situated at his right hand. A preliminary practice was given until the subject felt confident that he could do what was required.

The threshold of fusion was determined by the following modification of the constant method. A pair of flashes with a long inter-flash interval (I.F.I.) was presented; if the subject pressed twice, flashes with a short I.F.I. were presented; if the subject then pressed once the procedure was repeated with decreasing range until two I.F.I.'s a millisecond apart were reached. Pairs of flashes at these I.F.I.'s were then twice repeated and if the responses were maintained the I.F.I. at which the subject pressed once was taken as the threshold.

The following sequence is from a typical record: 100–2, 50–1, 90–2, 60–1, 70–1, 80–2, 70–1, 75–1, 78–2, 75–1, 76–1, 76–2, 75–1, 75–1, 76–2, 75–1, 76–2.

The use of the term threshold may be objected to in that it is not the point at which 50/50 'one' or 'two' responses were given. To reach this 'classical' type of threshold apparatus of greater accuracy and full use of the constant method would be required. In the present instance this degree of accuracy is not required. A threshold I.F.I. measured to the nearest millisecond gives about 100 scale divisions from 40 to 140 msecs. with which to measure cortical activation. The reliability of thresholds measured in this way is 0·96 based on the data from 38 schizophrenics and 20 normals with thresholds taken on two occasions four minutes apart.

(*B*) *Measurement of reaction time*

The stimulus used was one of the 5 msec. 50 ft. lambert flashes employed in the threshold measurement. A millisecond dekatron timer was started by the closing of relay contacts in a circuit associated with the trigger circuit which presented pulses to the C.R.T. Constant errors in the circuitry were measured and deducted from the value of RT shown by the timer. The timer was stopped by the pressure of a button on the right arm of the subject's chair. Pressure rather than release was used as it had been found in the past that it was more difficult to persuade deteriorated schizophrenics to release a button than to press it. This was particularly so when they had been pressing the same button to signal whether they saw one or two flashes of light in the threshold measurements.

(*C*) *Measurement of skin potential level* (*second experiment only*)

Full details of the method of measurement are given in Venables & Sayer (1963). In this instance Ag/AgCl electrodes were used with 0·5 g./100 ml. KCl as the electrolyte. Electrodes were placed on the hypothenar eminence of the left hand and over an abraded portion of the inner surface of the left forearm. Abrasion by sandpaper achieved a source resistance in all subjects of less than 100 kΩ. With this source resistance and selected electrodes skin potentials were measured to the nearest 0·1 mV. Potential artifacts due to the use of non-isothermal electrodes were automatically compensated.

(*D*) *Measurement of card sorting time* (*third experiment*)

Thirty-six cards were sorted into two piles according to whether they had printed on them either a *B* or a *Z* appearing among eight other letters. The data used in this study were the fastest times (in seconds) achieved by each subject in four consecutive sorts. The procedure was similar to that which had been used in a previous experiment (Venables, 1963*b*).

(*E*) *Measurement of verbal behaviour* (*fourth experiment*)

Single word responses to each of twenty words from the Kent & Rosanoff (1910) word list were recorded. The twenty stimulus words were selected from the 100 in the original list, to have responses of which the probability (Russell & Jenkins (1954) norms) of the most likely word was not greater than 0·5 and so that as far as possible there was a linear decline in probability with the next most probable words. The average probabilities of the 1st to 8th most likely words were in order: 0·32, 0·19, 0·10, 0·06, 0·04, 0·03, 0·02, 0·02.

Procedure (*Experiments* 1 *and* 2)

Subjects were seated in front of the apparatus in darkness (except for the fixation patch) for 3 min. before the determination of two flash thresholds. Following threshold determination 21 RT's were measured. Stimuli were presented without warning at intervals which varied randomly between 7·0 and 8·5 sec. (It is difficult to decide in an experiment like the present one whether or not to use a forewarning. If a forewarning in another modality is used the effects of inter-cortical area or inter-modality activation have to be considered. If, on the other hand, a forewarning in the same modality is used there is a high likelihood of confusion between stimulus and warning with deteriorated schizophrenics. An inter-stimulus interval of around 8 sec. is probably sufficient to allow desynchronization of the alpha from the previous stimulus and response to terminate before the presentation of the next stimulus.)

In Expt. 2, the electrodes for skin potential measurement were affixed to the left hand and arm at the start of the experiment and a continuous recording made throughout the period.

In Expt. 3 two-flash thresholds were measured immediately before the first card sorting. About 1 min. elapsed between the end of threshold measurement and the start of the first card sorting; the time between each sorting was about 1 min.

In Expt. 4 two-flash thresholds were measured at the start of the experiment; responses to the twenty stimulus words which were spoken by the experimenter were obtained immediately after.

Subjects

All schizophrenics tested in these experiments had been in hospital for 2 years or more and could thus be classified as chronic patients. All were within the age range 25–59 years. There was definite evidence for all patients that they had had delusions, schizophrenic thought disorder, or catatonic motor disturbances, at some time during their illness. No patient had a history of epilepsy, organic cerebral disease, or was leucotomized.

Normal subjects were drawn from the nursing, medical, and artisan staffs of the hospital in which the study was carried out.

Experiment 1. Fifty-six patients took part; forty-two of these were receiving no medication at the time of this experiment. Fourteen were receiving various doses of Stelazine (trifluoperazine). The patients were seen on one occasion only.

Experiment 2. Twenty patients took part on three occasions 1 month apart. None were receiving drugs on any occasion or between occasions. Twenty normal controls were also tested on the first occasion. Only seventeen were eventually available to provide data on all occasions.

Experiment 3. Thirty-two patients, none of whom were on drugs, were tested.

Experiment 4. Sixty-three schizophrenic patients and sixteen normal subjects were tested. Twenty-four of the patients were receiving various dosages of chlorpromazine; the remainder were unmedicated.

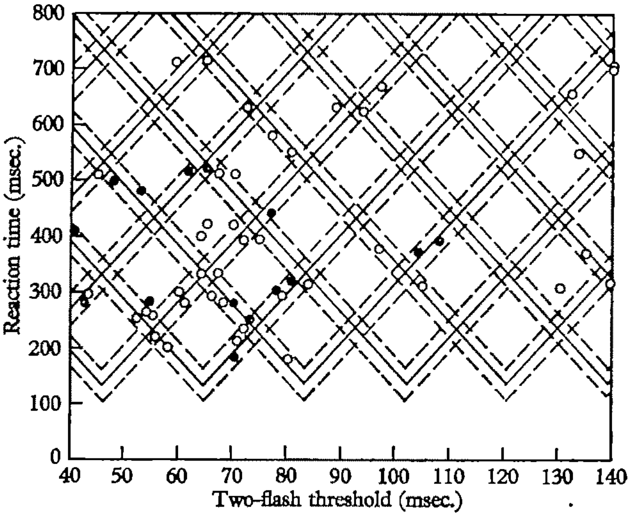


Fig. 1. Reaction times and two-flash threshold of 55 schizophrenic subjects with arbitrarily fitted multiple V pattern. Open circles, patients not on drugs; filled circles, patients on Stelazine (Expt. 1).

III. RESULTS

Experiment 1

The median of the 21 RT's given by each subject was plotted against the value of his threshold of paired flashes. (Twenty-one RT's were recorded so as to obtain an actual RT as the median. An even number of RT's would have resulted in a calculated median.) The resulting plot for 55 of the 56 subjects is given in Fig. 1. The remaining subject had a RT greater than 800 msec. From preliminary experiments relating skin potential (as an activation index) and RT it had become evident that a single U-shaped function could not encompass the data. Evidence of multiple U's in this early work suggested that this pattern of relationship might be of use in describing the present data. The X-axis spacing of the V's in Fig. 1 was suggested by converting potential indices from the earlier experiments into thresholds by means of regression lines given in Venables (1963*a*). Multiple V's rather than U's were found to achieve the least ambiguous fit to the data. The slopes of the V's were decided by trial and error.

In order to decide whether the resulting pattern was a significant fit to the data, limits were drawn on either side of the V lines so that there were equal areas of 'line' and 'space'. Reading directly from the plot, data from 47 of the subjects fell on the 'lines' and that from eight on the 'spaces'. By extrapolation the data from the single subject outside the limits of Fig. 1 fell on a 'line'.

A χ^2 test of this distribution against equal expected frequencies in 'line' and 'space' areas gave a value of 28·6 ($P < 0{\cdot}001$).

It has thus been shown that the data relating simple reaction time and a measure of cortical activation can be subsumed under a multiple V pattern. To show that this pattern is not one of many which might be chosen to fit the data but that it has some meaningful status Expt. 2 was carried out.

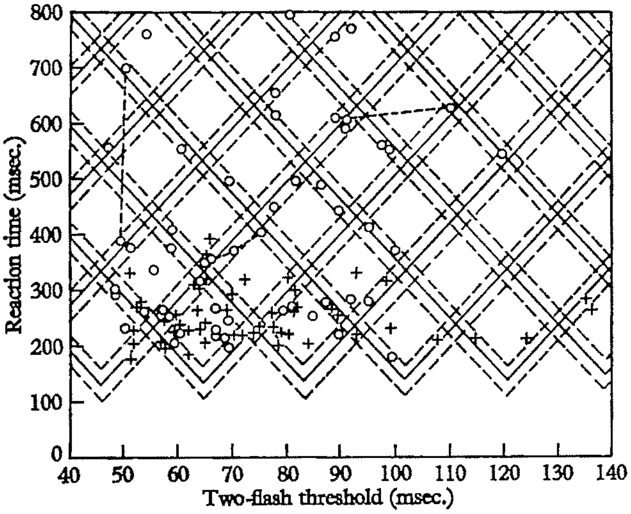


Fig. 2. Reaction times and two-flash thresholds of 20 schizophrenic and 17 normal subjects on three occasions. Parameters of pattern identical with those of Expt. 1. Open circles, schizophrenics; crosses, normals (Expt. 2).

Experiment 2

Data relating threshold of fusion of paired flashes and median RT for 20 schizophrenic and 17 normal subjects on three occasions are plotted in Fig. 2. The parameters of the multiple V pattern in this figure are those of the first experiment. Forty-two out of the 60 points given by the schizophrenic subjects fell on 'lines'; 18 fell on 'spaces'; of the 51 points given by normal subjects 29 fell on 'lines' and 22 fell on 'spaces'.

Testing these data against equal expected frequencies of points in 'lines' and 'spaces' the schizophrenic data show a significant fit to the multiple V pattern ($\chi^2 = 9{\cdot}6$, $P < 0{\cdot}01$). The data from normal subjects do not show a significant fit to the proposed pattern ($\chi^2 = 0{\cdot}98$, NS). The data are shown in an alternative form in Table 1.

Table 1. *Goodness of fit of data from Expt. 2 to proposed multiple* U *pattern at different RT levels*

	Normal		Schizophrenic		Total		
Interval (msec.)	On line	In space	On line	In space	On line	In space	% on line
180–189	1	0	1	0	2	0	100
190–199	1	0	1	0	2	0	100
200–109	5	2	1	1	6	3	66
210–219	3	2	1	0	4	2	66
220–229	2	7	0	3	2	10	17
230–239	0	3	0	1	0	4	0
240–249	1	0	1	1	2	1	66
250–259	4	3	2	1	6	4	60
260–269	1	1	3	2	4	3	57
270–279	1	0	1	1	2	1	66
280–289	3	0	0	0	3	0	100
290–299	1	0	1	0	2	0	100
300–349	5	3	3	1	8	4	66
350–399	1	1	6	2	7	3	70
400–449	.	.	5	0	5	0	100
450–499	.	.	3	0	3	0	100
500–549	.	.	2	0	2	0	100
550–599	.	.	4	0	4	0	100
600–649	.	.	4	0	4	0	100
650–699	.	.	2	0	2	0	100
700–749	.	.	0	0	0	0	.
750–799	.	.	0	4	0	4	.
800+	.	.	1	1	1	1	50
Totals	29	22	42	18	71	40	.

The goodness of fit to the multiple V pattern is examined at different levels of RT. The RT intervals are graded in 10 msec. steps from 180 to 300 msec. and thereafter in 50 msec. steps.

From this table it can be seen that the fit to the multiple V pattern is good for RT's above 280 and below 750 msec. ($\chi^2 = 23{\cdot}2$, $P < 0{\cdot}001$ for combined normal and schizophrenic data). RT's between 200 and 280 do not appear to fit the pattern, but the four RT's below 200 msec. do show a clear fit.

The data in Fig. 2 are from 37 individuals on three occasions; they were examined to see whether there was a tendency for individual data subjects to remain on one V from occasion to occasion or to change from V to V. No consistent tendency towards either of these alternatives was seen. The data on three occasions for three subjects are shown linked together as examples of what can happen. (This is not done for all subjects in order to maintain a readable figure.) It can be seen that while the data for one subject are more or less maintained on a single V those for another remain closely on the same threshold and show an alteration of RT. The third example shows a maintenance of RT level and a change of threshold between occasions.

It could be suggested that the multiplicity of equally spaced V-shaped relationships could be due to some sort of cyclical artifact in the measurement of fusion of paired flashes. To counter this objection reaction time was related to skin potential measured at the end of the threshold determination.

It has been shown that in schizophrenics these measures of autonomic and cortical activation level are closely related ($r = -0{\cdot}72, P < 0{\cdot}001$). Regression lines relating the two measures are available (Venables, 1963*a*). It has also been suggested that while skin potential level may possibly be considered as an index of the activity of the reticular system, the cortical activity level, as measured by threshold of fusion of paired flashes, is a function of the interaction of activating and regulatory functions of lower centres on the cortex. It has been proposed that the high correlation between autonomic and cortical activity levels in the schizophrenic is due to the absence of regulatory activity. The presence of this latter function in normals leads to a negative relationship between cortex and autonomic system.

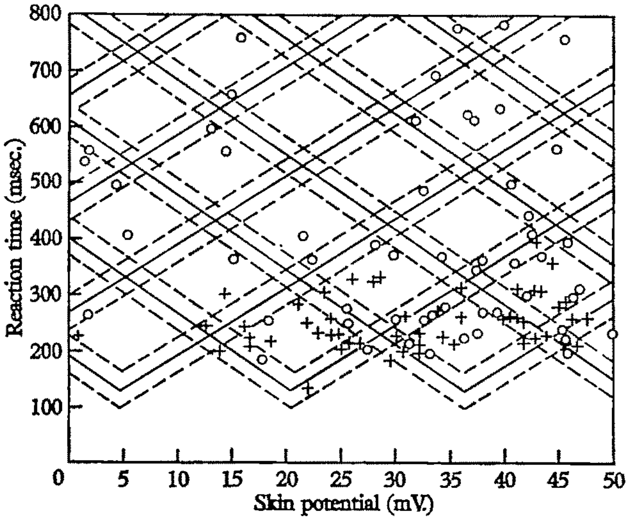


Fig. 3. Reaction times and skin potentials of subjects in Expt. 2. Multiple V pattern derived directly from that in Figs. 1 and 2, see text. Open circles, schizophrenics; crosses normals (Expt. 2).

Of the available regression lines relating two-flash threshold and potential, that which gives threshold in terms of potential for schizophrenics appears to be the correct one to use (Th. $= -1{\cdot}24$ Pot. $+ 109{\cdot}2$). It is presumed 'uncontaminated' by the regulatory activity shown in the normal and relates cortical activity in terms of a, so to speak, prior function of a lower centre. The equation in the form

$$\text{Potential} = \frac{\text{Threshold} - 109{\cdot}2}{-1{\cdot}24}$$

was therefore used to redraw the multiple V-shaped pattern in terms of potential and RT.* This is shown in Fig. 3. Boundary limits to give equal 'lines' and 'spaces' were drawn as in Figs. 1 and 2.

---

* As a check on this argument the multiple V pattern was drawn using the regression line of potential on threshold, calculated from the same data, Pot. $= -0{\cdot}45$ Th. $+ 61{\cdot}8$. The resulting multiple V pattern did not fit the data.

The data relating RT and skin potential are plotted in Fig. 3. Out of the 60 points derived from the schizophrenic data 40 fall in 'lines' and 20 on 'spaces' ($\chi^2 = 6{\cdot}6$, $P < 0{\cdot}01$). The data for normal subjects do not fit the pattern, nor does there appear to be any improvement of fit for RT's above 280 msec. as seen with the threshold data. Calculation of a fresh multiple V pattern for normals in terms of the regression line relating threshold and potential for these subjects in the resting state (Venables, 1963*a*) resulted in a slight but not significant improvement.

Experiment 3

To determine how far the pattern of results obtained with reaction time was capable of being generalized to other tasks in which speed of performance is a relevant measure, the relationship between card-sorting time and two-flash threshold was examined. Time per card was calculated as the mean time to make 36 two-choice sortings. This value was plotted for each subject against his two-flash threshold measured immediately before the commencement of card sorting. Using the same parameters as those developed in Expts. 1 and 2 a fit to the present data could not be obtained. However, choosing V's of different slope with the same values of two-flash threshold for the points of intersection of the V's as those found for the RT experiments enabled a significant fit to be made to the data. As before, equal areas of 'line' and 'space' were drawn; data from 24 of the 32 subjects fell on 'lines' ($\chi^2 = 8{\cdot}0$, $P < 0{\cdot}01$). The slope of the best fitting V's in this experiment was 73·5 msec. of sorting time per msec. of two-flash threshold; a figure which is almost exactly seven times that for the RT data.

Experiment 4

The three previous experiments have been concerned with the relationship between a measure of cortical activation and a measure of efficiency of performance expressed in terms of time. Efficiency of performance can, however, be expressed in other ways. If the use of language can be considered as performance, the coherence with which it is used can perhaps be considered as a measure of efficiency. A range of degrees of incoherence of speech may be found in a population of schizophrenics. Coherence of speech may be thought of as the extent to which it obeys rules set by grammar, syntax and meaning. Thus the appearance of any particular word in a coherent sentence is governed by the words which went before and to some extent by the words which follow it. The choice of it rather than a number of other words which might take its place may be expressed as a probability that it will be chosen out of a finite population of other words. Confining the scope of the inquiry to that of meaning, the probability of a single word response in a word association experiment may be expressed in terms of the number of other words that have been given as responses to the stimulus words. Osgood (1957) has suggested ways in which the relative hierarchies among a set of probabilities of verbal response alternatives may be changed under different motivating or arousing states. His explanation provides the starting-point for this experiment.

Russell & Jenkins (1954) provide response norms for the 100 words of the Kent–Rosanoff (1910) word list. Responses were obtained for each subject for each of twenty words from this list. The mean probability of each subject's responses was calculated by taking the mean of the summed square roots of the probabilities of each response word. (For small values of P the more correct method of taking the mean of the summed arc sine transformations of P is approximated by using $\sqrt{P}$.) The data thus obtained were plotted against the two-flash threshold measured immediately before presentation of the word list.

Using the same points for intersection of the V's as in the previous experiments a fit to the data could be obtained. For those schizophrenics not on drugs the data for 29 fell on 'lines' while that for 10 fell on 'spaces' ($\chi^2 = 9{\cdot}48$, $P < 0{\cdot}01$). For those schizophrenics on drugs the data for 16 fell on 'lines' and 8 on 'spaces' ($\chi^2 = 2{\cdot}66$, NS). For the normals the figures were 11 on 'lines' and 5 on 'spaces' ($\chi^2 = 2{\cdot}25$, NS). The data are shown graphically in Fig. 4, which is plotted in the inverse direction to previous figures with coherence increasing up the Y-axis.

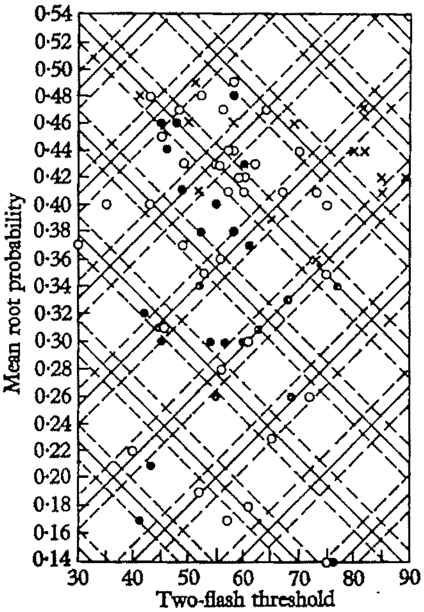


Fig. 4. Mean root probabilities of responses to 20 words from the Kent–Rosanoff list and two-flash thresholds for 63 patients and 16 normal subjects in Expt. 4. Open circles, schizophrenics not on drugs; closed circles, schizophrenics on chlorpromazine; crosses, normals.

As found in Expt. 2 it seems that a less good fit to the multiple V pattern is obtained with near optimal data; thus that for normals is not a significant match to the pattern. The lack of significant fit with the data from those patients receiving chlorpromazine is due largely to slight deviations from the pattern fitted to the non-drugged schizophrenics rather than to any gross deviations from that shown by the latter patients. The data provided by this experiment are presented very tentatively as it is felt that although logically reached the decision that led to the use of the square root transformation of the original probabilities has resulted in data which appear to be linearly related in a multiple V fashion to the two-flash data by some degree of serendipity. In an effort to provide a further check on the status of the patterning of the data presented, a set of 50 pairs of data was generated from adjacent columns of a table of random numbers, a restriction being imposed to produce data with the same range as that provided in practice. On the basis of the identical pattern which had been found to fit the verbal data 23 items of the artificial data fell on 'lines' and 28 on 'spaces'; the best attempt to fit the artificial data by changing the slopes of the V's while maintaining points of the V's at the same threshold values resulted in 31 items of data falling on 'lines' and 19 on 'spaces' ($\chi^2 = 2{\cdot}88$, NS). It would thus appear that the verbal data presented have some claim to be considered as having status in reality.

IV. DISCUSSION

The results enable the following statements to be made. The majority of the data relating RT and an index of cortical activity fit a multiple V-shaped pattern with equally spaced V's. There is a band of RT's between 200 and 280 msec. where the fit to the pattern is not maintained. Individual subjects do not necessarily provide data confined to a single V. The data from schizophrenics fit the multiple V pattern transcribed in terms of skin potential. In normals, owing no doubt to interaction of a regulatory influence between lower centres and cortex which disturbs the clear relation between activation at these two levels, the pattern of potential and RT directly transcribed from the threshold data does not fit. This patterning of results is not confined to reaction time but is also found in card sorting and it would appear even when the performance is not measured in terms of speed. However, in this latter instance as the speed of making a response is probably related to the range of probabilities of alternative responses which are considered, the verbal performance described may be a secondary function of speed.

The simulation of these results by means of an *ad hoc* pseudo-physiological model could be attempted. However, the likelihood that a model artificially constructed to encompass the results would have any physiological relevance seems doubtful.

Explanations of the single inverted U-shaped function relating performance and activation level have been attempted by various writers, for example Hebb (1955), Malmo (1959) and Welford (1962). The improvement in performance with increasing activation level up to an optimum point is explained in various related ways. Hebb talks of enhanced synaptic transmission; Welford an increase in system sensitivity and responsiveness, and Malmo the facilitation of the circulation of neural impulses in closed chains. To achieve the magnitude of changes in RT found in practice it seems unlikely that changes in synaptic delay or even in the number of parallel neurones firing can be responsible. Rather is it necessary to invoke varying numbers of successive delays caused by the series accumulation of circulation times of reverberatory circuits. Enhanced sensitivity of the system with increased facilitation from the activation system could thus be expected to decrease response time by the dropping out of delay elements formerly required temporally to summate impulses in order to overcome high synaptic thresholds.

The worsening of performance due to supra-optimal arousal levels has also various explanations. Hebb suggested the facilitation of irrelevant responses, Malmo the growth of high synaptic thresholds in circulatory chains due to excessive bombardment, and Welford the random firing of neurones leading to noise in the system and a decrease in channel capacity.

Although differently expressed it is possible that the apparent contrast between these suggestions arises because they are on different levels of explanation.

If it can be accepted that the magnitude of differences in RT found in the present experiments is such as to involve the comparatively large delays which are produced in closed neuronal chains, it is not a very remote step to see how multiple U- or V-shaped patterns can be produced. The momentary sensitivity of motor neurones may demand for firing at any one time the summation of one more or one less impulse than previously. The reaction time at any one activation level may therefore rise or fall by the reverberation time of neuronal chain. Examination of Figs. 1 and 2 suggests that this time (the vertical distance between V's) is of the order of 200 msec. This is in contrast to a previous finding (Venables, 1960) suggesting 100 msec. as the time of reverberation. While not required for an economical summary of the data a multiple V pattern with V's at 100 msec. vertical separation would undoubtedly encompass some of the points which are in 'spaces' in the present pattern. The present data do not therefore rule out a 100 msec. periodicity.

As a rather extreme speculation it is suggested that a 33 msec. 'quantum' of reverberation time as suggested by Augenstine (1958) is the basic building brick of the system. This would give a stepwise increase or decrease in RT with increase or decrease in activation level that is sufficiently smooth to be in accord with the data. It is further suggested that relatively smooth V or U shapes resulting from 33 msec. quanta can be switched in relation to activation level by 100 or 200 msec. circuits or by combinations of three or six 33 msec. circuits acting together to give the appearance of multiple V's.

It is suggested that the relative independence of activation level of RT's having values between 200 and 280 msec. is evidence of the optimal operation of the system without circulatory delays. This is also suggested by the data in Expt. 4 where $\sqrt{P}$ is greater than about 0·42. It is only when the system is other than optimally sensitive that the relevance of the arousal system in the determination of the operation of delay circuits becomes manifest.

However, this is, as has been said, extreme speculation; further work is required to establish the validity of the general pattern and to extend the work by determining what features of the experimental situation change the parameters of the pattern. The value of results would seem to lie in the light they throw on why it has been difficult heretofore to discover simple relations between performance and arousal.

This work was carried out at Netherne Hospital, Coulsdon, Surrey, by kind permission of the Medical Superintendent, Dr R. K. Freudenberg, and with the help of Mr P. Hill.

REFERENCES

AUGENSTINE, L. G. (1958). Human performance in information transmission. Pt. VI. Evidence of periodicity in information processing. *Report R* 75. Control Systems Laboratory, University of Illinois.

CALLAWAY, E. (1962). Factors influencing the relationship between alpha activity and visual reaction time. *E.E.G. Clin. Neurophysiol.* **14**, 674–82.

COURTS, F. A. (1942). Relations between muscular tension and performance. *Psychol. Bull.* **39**, 347–67.

DUFFY, E. (1951). The concept of energy mobilization. *Psychol. Rev.* **58**, 30–40.

FEDIO, P. M., MIRSKY, A. F., SMITH, W. J. & PARRY, D. (1961). Reaction time and E.E.G. activation in normal and schizophrenic subjects. *E.E.G. Clin. Neurophysiol.* **13**, 923–6.

FREEMAN, G. L. (1940). The relationship between performance level and bodily activity level. *J. Exp. Psychol.* **26**, 602–8.

FREEMAN, G. L. (1948). *The Energetics of Human Behavior.* New York: Cornell University Press.

FURTH, H. G. & TERRY, R. A. (1961). Autonomic responses and serial learning. *J. Comp. Phys. Psychol.* **54**, 139–142.

HEBB, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.* **62**, 243–54.

KENT, G. H. & ROSANOFF, A. J. (1910). A study of association in insanity. *Amer. J. Insan*, **67**, 37–126.

LANSING, R. W. (1957). Relation of brain and tremor rhythms to visual reaction time. *E.E.G. Clin. Neurophysiol.* **9**, 497–504.

LANSING, R. W., SCHWARTZ, E. & LINDSLEY, D. B. (1959). Reaction time and E.E.G. activation. *J. Exp. Psychol.* **58**, 1–10.

LINDSLEY, D. B. (1951). Emotion. In STEVENS, S. S. (ed.): *Handbook of Experimental Psychology.* New York: Wiley.

MALMO, R. B. (1959). Activation: a neuropsychological dimension. *Psychol. Rev.* **66**, 367–86.

O'HARE, J. J. (1954). The variability of auditory and visual reaction time with change in amplitude and phase of alpha rhythm. *Amer. Psychol.* **9**, 444.

OSGOOD, C. E. (1957). Motivational dynamics of language behaviour. Chapter in *Nebraska Symposium on Motivation* (JONES, M. R. (ed.)). Lincoln: University of Nebraska Press.

RUSSELL, W. A. & JENKINS, J. J. (1954). The complete Minnesota norms for responses to 100 words of the Kent–Rosanoff Word Association Test. *University of Minnesota Technical Report* no. 11.

SCHLOSBERG, H. & KLING, J. W. (1959). The relationship between 'tension' and efficiency. *Percept. Mot. Skills*, **9**, 395–7.

STENNETT, R. G. (1957). The relationship of performance level to level of arousal. *J. Exp. Psychol.* **54**, 54–61.

SURWILLO, W. W. (1963). The relationship of simple response time to brain wave frequency and the effects of age. *E.E.G. Clin. Neurophysiol.* **15**, 105–14.

VENABLES, P. H. (1960). Periodicity in reaction time. *Brit. J. Psychol.* **51**, 37–43.

VENABLES, P. H. (1963*a*). The relationship between level of skin potential and fusion of paired light flashes in schizophrenic and normal subjects. *J. Psychiat. Res.* **1**, 279–87.

VENABLES, P. H. (1963*b*). Selectivity of attention, withdrawal and cortical activation in chronic schizophrenics. *Arch. Gen. Psychiat.* **9**, 74–8.

VENABLES, P. H. & SAYER, E. (1963). On the measurement of skin potential. *Brit. J. Psychol.* **54**, 251–6.

WALSH, E. G. (1952). Visual reaction time and the alpha rhythm, an investigation of the scanning hypothesis. *J. Physiol.* **118**, 500–8.

WELFORD, A. T. (1962). Arousal, channel capacity and decision. *Nature, Lond.*, **194**, 365–6.

(*Manuscript received* 8 *May* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 219–225
Printed in Great Britain

THE SPIRAL MAZE

A PSYCHOMOTOR TEST WITH IMPLICATIONS FOR THE STUDY OF DELINQUENCY

BY H. B. GIBSON
Institute of Criminology, University of Cambridge

The writer discusses previous work on the use of psychomotor tests in the discrimination between normal subjects and subjects who are maladjusted or delinquent. A new test, the Spiral Maze, is described. Details of two experiments using the Spiral Maze are described: (*a*) contrasting a group of 22 young delinquent boys with 39 normal controls; and (*b*) testing 234 normal schoolboys aged 8–9 for whom teachers' ratings on 'naughtiness' were available.

It was found that: (i) the delinquent group were significantly different in their preference for speed at the expense of accuracy; (ii) there was a significant relationship between psychomotor scores and the categories of 'naughty', 'average' and 'good' which were derived from the teachers' behavioural ratings; (iii) although, as expected, 'naughty' boys predominated in the group which was quick and careless, this 'naughty' category also predominated in the slow and careless group.

The implications of these findings are discussed.

Psychomotor tests are concerned with measuring the speed, accuracy and style of subjects' movement in response to various stimuli. Used in personality research, they have the great advantage of having little obvious connexion with such matters as neuroticism and behaviour disorder, and hence they do not provoke attitudes which lead to the falsification of test responses. In general, they are simple to administer, and scorable by entirely objective means. What they lack is any very clear theoretical background which explains why such test responses in a psychological laboratory give valid indications of the subjects' past and future social behaviour.

Among early work, psychomotor performance was studied by Davis (1948) in the course of his cockpit study. Experiments with pilot trainees showed associations between psychomotor movement and both neurotic syndromes and accident-proneness. Errors in movement were measured in terms of their amplitude and duration; at one extreme there were subjects whose performance suffered from overactivity, and at the other extreme those who were too inert. Both extreme types were accident prone, and there was some evidence that the overactive type were subject to neurosis of a dysthymic type, whereas the inert subjects were of a type originally categorized by Davis as 'hysteric'.

Venables (1955) followed up some of this work with industrial employees and neurotic patients. An interesting aspect of Venables's research is that there was a small group of industrial employees who were subsequently discharged, largely for dishonesty, who gave psychomotor responses of a pattern significantly different from that of the other employees. This led Sylvia Anthony to repeat Venables's work. The advantage of her study (Anthony, 1960), is that it was predictive. She used R.A.F. candidates and she was able to follow up the records of these young men after they had been tested and to compare the records of those who were later on convicted of offences with those who were not. She found that the psychomotor performance of those who subsequently became offenders differed from the normal. Their performance was rather like that of the group whom Venables designated as quasi-hysterics, and Anthony interprets her results in terms of Davis's suggestion that psychomotor disturbance is due to anticipatory tension.

Any discussion of psychomotor tests must make mention of the Porteus Maze Test (Porteus, 1959). In its original form this was regarded simply as an intelligence test; in Porteus's terms it is regarded as testing the capacity to plan ahead. In 1923 Porteus called attention to the fact that although delinquents might score normal test quotients, their performance on the mazes appeared to be qualitatively different from that of normal subjects. It was not until 1941, working with Honzik, that he arrived at a system of allocating points for poor qualitative score ('Q'); a high Q score indicates that the subject has produced a careless, slovenly performance and has shown certain characteristic errors of execution, even though he may have solved the mazes quite successfully. Subsequent research by many different workers has confirmed the finding that delinquent populations obtain significantly higher Q scores. There have also been some studies, such as that by Foulds (1951), which show that psychoneurotic populations obtain abnormal Q scores.

Now the Q score is made up of several more or less distinct categories of error; one of these categories comprises errors caused by cutting corners, touching the sides of the alleyways and making slovenly wavy lines. This is the psychomotor component of the test, and the fact that delinquent populations obtain high Q scores is partly due to the fact that such people are significantly different in psychomotor style.

One of the drawbacks of the numerous studies which report a high Q score with the Porteus Maze among delinquent populations is that we do not know to what extent the fact that delinquents have been detected, arrested and incarcerated affects their test performance. However, Anthony's study, which was predictive, indicates that psychomotor style is related to an important personality variable, and is not simply a transitory artifact of the test situation. The study reported here falls into two parts; the first part concerns established criterion groups, and the second is predictive in that we must await the future social development of the experimental population before the validity of the test measures can be fully tested.

EXPERIMENT A

Subjects. The subjects were in two groups: (i) 39 schoolboys taken from different academic streams so that a wide scatter of intellectual ability was represented; (ii) 22 delinquent boys in a remand home. The schoolboys were aged 11–12 years (median 11·5), and the delinquents aged 12–13 years (median 12·75).

Test material and procedure. The Spiral Maze is printed on cards 10 in. square. It consists of a heavy black line printed in spiral form so that it borders an open pathway 135 cm. long. Obstacles, in the form of the letter O in heavy type, are scattered along the whole length of the spiral pathway.

The subject, tested individually, is instructed to trace a pencil line from the centre of the spiral to the outside without touching either the black lines at the side of the pathway or the obstacles. He is told to do this *as quickly as he can.* As soon as he begins the task a stopwatch is started, and every 15 sec. the experimenter reminds him, quite sharply, to go as quickly as he can. This repeated reminder is intended to induce time-stress. As soon as the pencil reaches the end of the pathway the watch is stopped.

From each record two scores are obtained: (*a*) the time taken, and (*b*) the number of errors made. Errors are scored in terms of the frequency with which the pencilled line comes into contact both with the border of the pathway and the obstacles: a weighting is given according to the degree of penetration at each contact.*

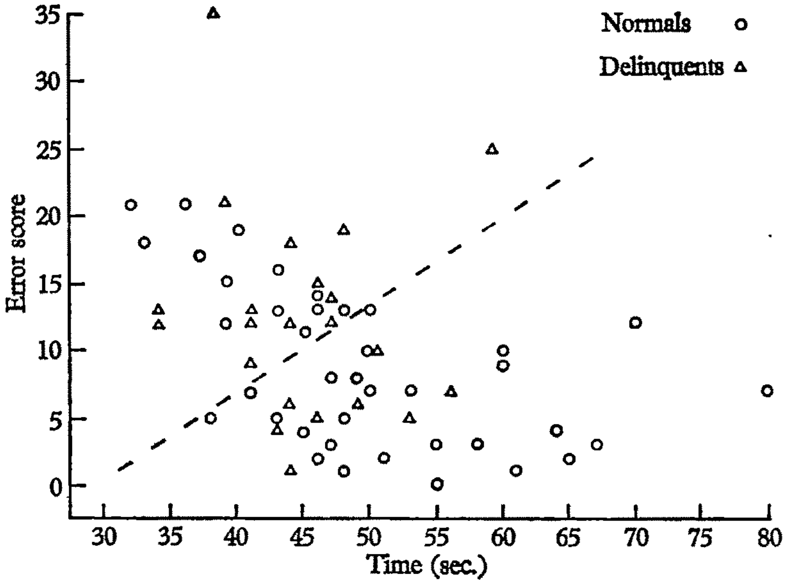


Fig. 1. Scatter plot of the scores of normal boys (circles) and delinquent boys (triangles) on the Spiral Maze.

Results

Fig. 1 shows the time and error scores of the two groups. It may be seen that there is a negative relationship between speed and accuracy ($r = -0{\cdot}42$, $P < 0{\cdot}001$) and that the delinquent group tends to be both quicker and less accurate than the normal group. The best discrimination between the two groups is represented by the diagonal line which has been drawn across the scatter plot. The value of χ^2 of the difference between the two groups, taking such a line as a cutting point, is 7·8 with 1 D.F., which is significant at the 0·01 level of confidence.

The difference between the scores of the two groups might conceivably be attributed to the fact that the normal group was mostly younger than the delinquent group. As no norms have yet been established concerning performance on the test at different age levels, the product moment coefficients between age and both time and error were computed for the normal group. Both values of r were quite insignificant ($-0{\cdot}15$ and $+0{\cdot}16$) and it is therefore unlikely that the observed difference was due to the slight difference in age.

EXPERIMENT B

Subjects. The subjects were 254 boys in the age range 8–10 years. They were pupils in primary schools and none of them had been charged in court.

---

* Supplies of test material together with full details of administration and scoring are available from the author.

Test material and procedure. The subjects were tested on the Spiral Maze as previously described, but in this experiment the test was applied immediately after the Porteus Maze series had been completed.

A schedule was completed for each boy by his class-teacher. This schedule consists of a series of ratings on general behaviour, freely adapted from that used by the National Survey (Douglas, 1956). The aspects of behaviour covered, and the provisional weightings allocated, are shown in Table 1.

Table 1. *Criteria on which teachers' ratings of 'naughtiness' are based*

	Maximum points
1. Truanting	5
2. Difficult to discipline	5
3. Not concerned to be a credit to his parents	5
4. Noticeably dirty and untidy	5
5. Lazy	4
6. Difficulty with peers due to aggressiveness, etc.	2
7. Seriously distractible	2
8. Not specially good at anything	1
9. Outstandingly bad in at least one activity	1
Maximum score	30

Results

The rating on behaviour was done by 17 school-teachers whose groups of boys ranged from 16 to 26 individuals, and the distributions of rating scores were divided within each group. Boys with scores in the uppermost quartile were labelled 'naughty', those with scores in the lowest quartile 'good', and the remainder were labelled 'average'.

In Fig. 2 the time and error scores of the 'naughty' and the 'good' boys are shown. Here, because of the skew of the distribution of time scores, time has been represented on a logarithmic scale.

In Table 2 the actual numbers of boys in the 'good', 'average' and 'naughty' categories are shown with respect to their positions on a scatter plot such as that given in Fig. 2, where median lines divide the plot into four quadrants. Analysis of variance of these data shows that the variance ratio *between categories* gives $F = 10{\cdot}18$ (D.F. 2 and 6) which is significant beyond the 0·05 level of confidence. It is thus demonstrated that the position of the boys' scores on the scatter plot is significantly related to their behavioural category as given by their teachers.

The above finding is clarified in Table 3. Here the numbers are represented as percentages of the behavioural categories, and the zones in the four quadrants are given appropriate titles.

DISCUSSION

It was demonstrated in Expt. A that scores on the Spiral Maze discriminated significantly between a normal and a delinquent group, the latter being more 'rash', that is, sacrificing accuracy to speed.

The object of Expt. B was to investigate whether there was any relationship between psychomotor abnormality and 'naughtiness', that is, bad behaviour and low social standards which come to the notice of school-teachers. Such 'naughtiness' is a well-known predictor of later delinquency.

One of the drawbacks about using school-teachers' ratings is that individual teachers have very different standards of strictness in rating behaviour. Thus one

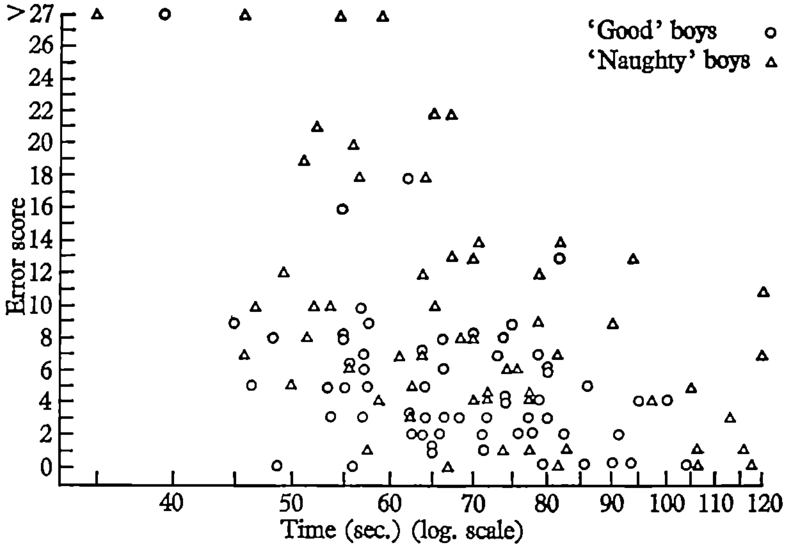


Fig. 2. Scatter plots of the scores of 'good' schoolboys (circles) and 'naughty' schoolboys (triangles) on the Spiral Maze.

Table 2. *Distribution of* 254 *normal schoolboys on a scatter plot of Time and Error scores on the Spiral Maze*

('Good', 'average' and 'naughty' are categories derived from teachers' ratings of general behaviour.)

CARELESS

		n		*n*	
	Good boys	15	Good boys	8	
	Average boys	42	Average boys	25	
	Naughty boys	24	Naughty boys	14	
QUICK					SLOW
	Good boys	18	Good boys	20	
	Average boys	22	Average boys	45	
	Naughty boys	6	Naughty boys	15	

ACCURATE

Table 3. *Percentage distributions of* 254 *normal schoolboys on a scatter plot of Time and Error scores on the Spiral Maze*

(The percentages are within each behavioural category.)

Teachers' behavioural category	Zones on the plot of Time and Error scores: Quick and accurate	Quick and careless	Slow and accurate	Slow and careless	Total
Good boys	29·5	24·6	32·8	13·1	100
Average boys	16·4	31·3	33·6	18·7	100
Naughty boys	10·2	40·7	25·4	23·7	100

teacher may rate 30 % of his class as 'difficult to discipline' and another teacher may say that *none* of his class are to be categorized thus. An early study of this problem (Hartshorne & May, 1928) indicates that one must accept this inter-rater variability as an inevitable source of error. Where the teachers are rating groups of a fair size, various adjustments can be made for inter-rater bias. In the present study this source of error was reduced by making the tripartite division *within each group*, as already explained.

Comparing the Spiral Maze scores obtained in Expts. A and B (Figs. 1 and 2) it may be seen at the level of inspection that the older group is quicker than the younger. It cannot be assumed that this difference is a function of age; the latter group were tested immediately after they had run the series of Porteus Mazes, and therefore they may have been slowed down by this pre-training, since the Porteus does not encourage speed. Norms on the Spiral Maze relating to age and conditions of testing have yet to be established, and the data reported here are complex.

It was expected that the scores of the 'naughty' boys would have differed from those of the 'good' boys, with respect to position on the scatter plot, in the same way that the delinquents' scores differed from those of the normal controls. Comparison of the two figures shows that this is not so. Table 3 shows that although in the quick-and-careless zone the scores of the 'naughty' boys predominate, as expected, they predominate also in the slow-and-careless zone. What characterizes the 'naughty' boys in all four zones is their relatively high error score. It may also be seen in Table 3 that the position of the 'average' boys is intermediate, quantitatively, between the 'good' and the 'naughty' in all zones except the slow-and-accurate zone, where the proportions of the 'average' and 'good' are almost identical.

The characteristics of the 'naughty' boy are not quite the same as those of the delinquent boy. While a proportion of these 'naughty' boys will probably become delinquent in a few years time, their present categorization is due entirely to impressions formed by their teacher. Future re-testing will show whether the slow and careless performers will tend to speed up, or to increase in accuracy when they are older. This follow-up will also show the predictive power for future delinquency of both the teachers' ratings and the measures derived from the Spiral Maze.

The crucial factor in this test is the relationship between speed and error. Running the Spiral Maze resembles some of the elements of a car-driving skill. The more the subject speeds, the more likely he is to collide with obstacles on the road and the kerb at the side. Very few subjects make no errors, but the effect of touching an obstacle is to make the subject slow down and adjust his speed to his degree of manual dexterity. This implies a proper sensitivity to feedback which comes from observing one's own errors. The subject is given no guidance as to whether speed or accuracy is the more important; this decision is left entirely to his own discretion. In a sense, therefore, the test is one of risk-taking in which the subject should ideally make no errors by proceeding slowly and cautiously, for the task is of elementary simplicity, but it has the added complication that the experimenter is repeatedly nagging at the subject to go as fast as he can.

In the Porteus test, the Q score errors result from 'inaccuracy' of two separate and sometimes unrelated kinds. Thus while entering an alleyway early on may be due to inadequate foresight, the cutting of corners is due not to lack of foresight but to crudity of psychomotor style. The measures derived from the Spiral Maze differ from the Porteus Q measures in two important respects. First, they are entirely psychomotor in origin; secondly, they are obtained under time stress, so that they are partly a function of how the subject reacts to such stress. This latter provision furthers an aim suggested by Foulds (1951), of 'the heightening of the temperamental and the lowering of the intellectual factors'. In Foulds's own modification of the Porteus test, timing was certainly introduced, but no attempt was made to induce time stress by repeated exhortations to work quickly. Foulds's differentiation between different categories of subject in terms of time measures might have been rather different had he introduced a form of administration which made his subjects acutely aware of the passage of time.

The Spiral Maze does, however, resemble the technique of Anthony, and of the other workers cited earlier, in respect of the influence of stress on psychomotor performance. The essential difference here is that while in the earlier studies the subject could not control the speed at which the stimuli were presented to him, with the Spiral Maze (as administered in the present two studies) the subject could work at whatever rate he chose, in spite of the experimenter's nagging. It would be premature to speculate on the probable speed and error scores which might be obtained on this new test by 'quasi-hysterics' and other such abnormal personalities, as no one has produced sufficient evidence for the nature of psychomotor responses to be understood very adequately. Fortunately, as the present study is a longitudinal one, these matters should be clarified by the future social development and the further re-testing of the subjects mentioned in Expt. B.

REFERENCES

ANTHONY, H. S. (1960). Anxiety as a function of psychomotor and social behaviour. *Brit. J. Psychol.* **51**, 141–52.

DAVIS, D. R. (1948). *Pilot Error: Some Laboratory Experiments.* H.M.S.O.

DOUGLAS, J. W. B. (1956). *School Teachers' Questionnaire.* National Survey of the Health and Development of Children, United Kingdom.

FOULDS, G. A. (1951). Temperamental differences in Maze performance. *Brit. J. Psychol.* **42,** 209–17.

HARTSHORNE, H. & MAY, M. A. (1928). *Studies in Deceit.* New York: Macmillan.

PORTEUS, S. D. (1959). *The Maze Test and Clinical Psychology.* Palo Alto, California: Pacific Books.

VENABLES, P. H. (1955). Changes in motor response with increase and decrease in task difficulty in normal industrial and psychiatric patient subjects. *Brit. J. Psychol.* **46,** 101–10.

(*Manuscript received* 21 *June* 1963)

Brit. J. Psychol. (1964), 55, 2, *pp.* 227–248
Printed in Great Britain

PUBLICATIONS RECENTLY RECEIVED

Fundamental Physical Theory and the Concept of Consciousness. By JOHAN H. GREIDANUS. Oxford: Pergamon. 1961. Pp. xii + 159. 30*s*.

This book is a philosophical treatment of consciousness written by a physicist, the author being Chief Engineer of the Royal Netherlands Fokker Aircraft Factory. He is concerned to give to consciousness, which he regards as an extra-physical concept, an operational status. His treatment of consciousness is thus in line with modern physics' treatment of matter, where it is functional relations that are investigated and described, the essential nature of matter being left unanswered and accepted as unanswerable.

The status of consciousness in present-day psychology is far from clear, recent text-books taking little notice of it. No reference is given in the index of Woodworth (1938), nor of Woodworth & Schlosberg (1955), while Osgood (1953) gives but one effective reference, and there the discussion is limited to neurology. Indeed, where the word 'consciousness' is used it seems to have changed its meaning from experience to behaviour and supposed neural mechanisms underlying behaviour. William James, writing in 1890 (the *Principles*) gives Consciousness a dozen references. His celebrated Stream of Thought chapter starts with the words: 'We may begin our study of the mind from the inside.' He goes on to take the extreme academic view that it is not sensation which is basic but rather thought, saying that: 'the only thing which psychology has a right to postulate at the outset is the fact of thinking itself, and that must first be taken up and analysed. If sensations then prove to be amongst the elements of thinking, we shall be no worse off as respects them than if we had taken them for granted at the start.' This reflects Descartes's earlier position, but it is difficult to defend today with the introduction of machines that solve problems, for we do not want to say that they are conscious though we might wish to say they think.

The present position is curious. Philosophers have traditionally taken thought as primary, while engineers and physicists favour matter; but it is engineers who have made the most dramatic contribution to problem solving, by devising mechanisms which solve problems, and as a result thinking itself looks more mechanical and less concerned with consciousness. But most curious, the development of logical and calculating machines (from Venn and Babbage) was greatly helped by the work of logicians developing propositional and class logic as an antidote to metaphysics (Frege, Boole and Russell). The analysis of thinking into small unitary steps, described by symbolic transformations represented in a logical symbolism, or by discrete mechanical processes, has succeeded in divorcing thinking from consciousness so that now psychological terms such as 'insight' which were thought to be explanatory look metaphysical. William James reads most oddly, and the work of experimental psychology as a whole seems irrelevant when we consider consciousness. Indeed we are forced back to metaphysics—or physics? In Wittgenstein we find the philosopher concerned with consciousness in its original sense but greatly influenced by the mechanical rigour of propositional logic, which indeed he did much to develop with his Truth Tables. In the *Tractatus* (1922) he wrote:

> 3·1 In the proposition the thought is expressed perceptibly through the senses.
> 3·2 In the propositions thoughts can be so expressed that to the objects of the thought correspond the elements of the propositional sign.
> 3·201 These elements I call "simple signs" and the proposition "completely analysed".
> 3·202. The simple signs employed in propositions are called names.
> 3·302. The name means the object. The object is its meaning.'

We find Wittgenstein mulling over the relation between thought and consciousness in his last writings: *Philosophical Investigations*, XI: 'Was it *seeing* or was it thought?—Do not try to analyse your own inner experience.'

Consciousness has faded from the study of mind to be discussed now more by physicists than by psychologists.

What has Greidanus got to tell us? He considers (p. 101) that consciousness has no localization in the nervous system, even though neurones are supposed necessary for consciousness. He does not think (p. 103) that consciousness need be compartmentalized into individuals, but rather that 'individuation has no more clarity and justification than has a condensation area of particularly dense specific physical interaction in space-time'. In discussing this, he suggests that people involved in conversation (or love) are not two separate consciousnesses, because molecular neuronal interaction comprises an airborne code link between them, while other links may: 'lead up to Cirius or to a nebula in the furthest depth of the universe'. The author by turns speculates on 'bioconsciousness' and 'cosmical consciousness', but his speculations tend to be so general that we often lose him; though without entirely losing respect for his endeavour. He disarms us with such phrases as: 'almost all considerations of this book are sufficiently vague to be largely independent of the exact wording of the assumptions'. This comment is less than fair, for he does try to write clearly where almost any set of words evokes in the reader countless nonsensical dicta by earlier writers. In his conclusion (p. 126) he disarms us again: 'In essence consciousness is unique'.

But, sadly, Wittgenstein should have the last word: *thereof one cannot speak, thereof one must be silent.*

R. L. GREGORY

A Theory of Cognitive Dissonance. By LEON FESTINGER. London: Tavistock. 1962. Pp. xii+291. 35*s*.

Deterrents and Reinforcement: The Psychology of Insufficient Reward. By DOUGLAS H. LAWRENCE and LEON FESTINGER. London: Tavistock 1962. Pp. vi+180. 32*s*.

Festinger's book, previously reviewed in this *Journal* (1958, **49**, 258–9), has come to be accepted as an important contribution to social psychological theory since its appearance in the United States in 1957. The present edition, which does not differ from the original text, marks the first publication of the book in Great Britain. The second monograph, with which this review is concerned, has appeared more recently and represents a forthright attempt to extend Festinger's theory of dissonance to account for the extinction of instrumental responses in animals. As a new form of cognitive-expectancy theory, it may prove to be influential in this area.

A persistent problem for theories of instrumental learning concerns the effects of partial reinforcement on extinction since it is a well-established fact that periodic omission of reward during training prolongs extinction when reinforcement finally is withheld. Traditionally, the phenomenon has been termed the 'partial reinforcement effect' but, as Lawrence and Festinger's experiments show, it is also true that increasing the effort an animal expends to reach reward increases resistance to extinction; delay of reward during acquisition exerts a similar effect.

The explanation proposed by the authors can be summarized in their own words:

> 'If an organism continues to engage in an activity while possessing information that, considered alone, would lead it to discontinue the activity, it will develop some extra attraction for the activity or its consequences in order to give itself additional justification for continuing to engage in the behaviour' (p. 156).

Such a statement poses a number of conceptual questions which the authors consider in turn. First, how do we know when an organism possesses information that would lead it to discontinue an activity? Lawrence and Festinger suggest that any factor inducing hesitation, apart from the transient effects of novel stimuli, may be said to represent information dissonant with performance of an instrumental response. Several experiments are described which demonstrate that partial reward, delay of reward, and effort expenditure constitute deterrents of this kind. Secondly, what variables affect the magnitude of dissonance? Here, it is assumed that as the magnitude of a deterrent increases, so the amount of dissonance induced by continuing to perform the response increases. It is further alleged that as dissonance increases, the greater will be an animal's attempts to reduce it and, therefore, the greater will be the development of 'extra attractions' in the situation. These added attractions are believed to arise when a test situation affords the animal an opportunity to satisfy other subordinate motives. Finally, then, how can we measure the development of 'extra attractions'? At this point, Lawrence and Festinger suggest that if an 'extra attraction' has developed in a given situation, it should act as an incentive. If it could be shown, in the absence of any extrinsic motivation, that an animal will learn a new action in order to enter the original situation, this would presumably constitute acceptable evidence for such an incentive.

To test some of the implications of their theory, Lawrence and Festinger conducted a series of sixteen experiments, the results of which tend to confirm the analysis they present. The central findings can be stated as follows:

(*a*) Resistance to extinction after partial reinforcement is not primarily a function of the ratio of rewarded to non-rewarded trials (Experiment 5); rather, it is a function of the total number of non-rewarded trials. However, non-rewarded experiences *per se* are not regarded as crucial. If rats are placed in the goal-box on non-rewarded trials but are required to run along the alley on rewarded ones, extinction occurs more rapidly as compared with animals that are required to run on non-rewarded and rewarded trials alike (Experiment 6). The authors interpret this finding to mean that resistance to extinction is increased when such experiences are related directly to an expectation of reward, implying that dissonance occurs only when an animal has performed an action and 'then finds that this action is not justified by its consequences'.

(*b*) If an animal is never rewarded in a given place but only delayed there, it continues to run to that place during extinction longer than animals that have always been rewarded there or animals that have been partially-rewarded there (Experiment 7). Secondly, partial reward experiences in a specific goal-box prolong extinction regardless of whether or not the runway has been associated with partial reward (Experiment 4).

(*c*) If, before extinction trials are given, the animals' deprivation schedules are discontinued so that they are no longer hungry, the same or even greater differences in resistance to extinction are obtained than would be the case if the animals were hungry during extinction (Experiments 8–13). These results appear to support an implication of the theory, namely, that once 'extra attractions' have been acquired, a reduction of the dominant motivation should enhance their effects since such 'attractions' are assumed to depend on an opportunity to satisfy subordinate drives.

To conclude, it is probably wise to suspend judgement of Lawrence and Festinger's dissonance concept for it represents little more than a preliminary framework to account for the partial reinforcement effect at present. Even though the theory is not sufficiently explicit in certain respects, its particular value lies in the fact that it has promoted a series of new experiments. These will have to be considered by adherents of other theories in this area since these new data create specific difficulties for alternative concepts.

ROGER STRETCH

Displacement of Concepts. By DONALD A. SCHON. London: Tavistock. 1963. Pp. xiv+208. 30*s*.

This is a book about creative thought and the emergence of new theories, inventions and attitudes. It is a lively book, well illustrated by examples both from the history of ideas and from current industrial invention. Dr Schon, a former teacher of philosophy and now Director of Technical Services in the U.S. Department of Commerce, can call on an unusual range of experiences. In approach his account is observational and discursive, rather than experimental and demonstrative. But if Dr Schon does not always convince, he assuredly stimulates.

In its basic theme the book is perhaps less original than Dr Schon supposes. It propounds the view that new ideas arise neither mysteriously from the unknown, nor merely by mechanical rearrangements of old elements, but as a result of a process of emergence, the essential nature of which can be described as a 'displacement of concepts'. By this Dr Schon implies that the process is fundamentally one of metaphor or analogy. Metaphor applies as much to concepts themselves as to linguistic forms, and can be regarded as 'taking the familiar as a projective model for the unfamiliar'. There are two aspects of the matter—the 'radical' aspect, the element of novelty involved in new concepts, and the 'conservative' function, the retention of the patterns of the old beneath their new dresses.

This account is not so very different from Piaget's twofold processes of 'assimilation' and 'accommodation'. And before this the Herbartian theory of 'apperception', Spearman's doctrine of 'the education of correlates' and 'noegenesis', the concepts of the 'schema' and of 'learning sets', have all been getting at much the same thing, the tendency for new mental content to be grafted on to the old stock. However, Dr Schon treats this theme with a new freshness, and with a lively topicality in the choice of his instances.

L. S. HEARNSHAW

Affect, Imagery, Consciousness. By SILVAN S. TOMKINS. Vol I. *The Positive Affects.* New York: Springer. 1962; London: Tavistock. 1964. Pp. xiv + 522. 65*s*.

This is the first of four volumes; three were originally projected but a fourth has been deemed necessary to cope with the elaboration of the theoretical framework. The second volume has now appeared and the reviewer has been privileged to see drafts of large parts of the other two but this review will be confined to the first.

Reviewing just the first volume has proved an almost impossible task. A great many important revolutionary hypotheses are advanced, and these are richly interconnected and supported or elaborated with evidence from a wide array of disciplines. It would be possible in a review to seize upon one or two hypotheses as critical and comment accordingly. However, not only would this presuppose a high degree of prescience but it would miss the major feature of the work, that it is a massive re-thinking of psychological science. The challenge it poses to our accepted models of man (and of our near neighbours) cannot be met simply by re-thinking each part, yet once again, in terms of our current knowledge. Nothing metaphysical is intended in this statement; it is simply that enough of Tomkins's hypotheses are so well established that we would need new and different data to rebut his others and thus maintain the ones to which we are accustomed.

The main task for the reviewer seems to be that of drawing attention to characteristics of Tomkins's work as a general model. This does not, in this case, seem irrelevant. The growth of psychology is hindered by the lack of adequate general unifying conceptions. Most of our work is done in restricted problem areas, and development is very often a matter of sensing the general characteristics and limits of these restricted areas. General conceptions help in these matters. Hebb's neurological model, information theory and cybernetics have each helped some psychologists in this way. Tomkins has gone a very great distance further. One of the marked inadequacies of our general psychological conceptions of man has been that the models have been too impoverished to guide psychological practice—observed behaviour is too rich by far. From a deep knowledge of psychology, physiology, ethology and biology, Tomkins has created a model of appropriate richness. His hypothesis of the centrality of the affect system is a major breakthrough in the face of our traditional models narrowly based as they are on the primacy of drives. The drives, manifesting as they do the boundary conditions of a relatively stable, homeostatic internal environment, are shown to be low in urgency and high in specific informational content about 'where and when to do what'. In contrast the affects manifest and amplify states and change of states of the C.N.S. which are no more stable than one could expect for any organism that is mobile in an already complex environment. The affects are in consequence high in urgency but low in specific informational content. In combination the affects and drives define a system that is essentially open—the logically sufficient conditions for its behaviour may arise from within the system or in its environment. The relevance of this system for the adaptiveness of the organism lies in the fact that it matches in its complexity 'the receptor, analyser, storage and motor mechanisms within the organism and a broad spectrum of environmental opportunities, challenges and demands from without' (p. 112). The high degree of error, cognitive and motivational, that is involved in the initial relative independence of these many internal subsystems is postulated by Tomkins as the necessary price for a chance for aggregates of humans to survive. Thus, while Tomkins in this first volume places a good deal of stress on the evolutionary significance of the affects (shades of Darwin and of that remarkable work of Olaf Stapledon's, *Last and First Men*) this only serves to heighten the role of learning; as is patently obvious in the later volumes.

With guiding assumptions that are contextualist rather than organic Tomkins finds man far more complex and indeterminate than most of us would find comfortable. We can be grateful, however, for the openness of mind that these assumptions give to his wide-ranging search in elaboration and test of his theories. These searches yield further theories that may turn out to be no less significant for us than those that sent him on his journeys, and, because he is little tempted to turn all else into logical derivatives, may even survive them.

F. E. EMERY

Personality: Dynamics and Development. By Irving Sarnoff. London and New York: Wiley. 1962. Pp. xv + 572. 59*s.*

This useful book, written by the Professor of Psychology at New York University, is an attempt to give a coherent account of personality and its development by making use of both psychoanalytical and learning theory concepts. This, of course, is difficult to do; it is rather like trying to write a book in both French and German at the same time. In reading it through with one's mental sets suitably arranged for learning theory concepts, analytical terms like super-ego and oedipal complex come as something of a shock. Conversely, if one reads it with mental sets suitably arranged for psychoanalytical speculation, learning theory data impinge uneasily. What seems to be needed is a third language into which these two can be translated, steering clear, as far as possible, from technical jargon peculiarly associated with either approach. This may be asking for the impossible, so it may not be surprising that Prof. Sarnoff has not been able to achieve it. He says firmly in his preface that 'psychoanalytical theory runs through most of this book as the thread upon which concepts from other theories are woven'. He chooses this rather than learning theory as his basis because he thinks that 'the superstructure of many of our theories of learning rests upon the capacities and idiosyncrasies of the albino rat'. He certainly has a point there!

Prof. Sarnoff's opening chapter on 'The scientific approach to the study of personality' is stimulating, and he returns to this theme in his last chapter on 'Problems in the scientific study of personality', which is equally lively. The rest of the book is a well-organized account of motivation and ego development, mechanisms of defence, super-ego, and guilt, and personality development from childhood to old age.

This should provide an excellent text for undergraduate and post-graduate students, particularly those interested in the clinical field.

RALPH HETHERINGTON

Elements of Psycho-Analysis. By W. R. Bion. London: Heinemann. 1963. Pp. v + 110. 15*s.*

This small book is difficult to review briefly, first, because of the author's condensed style; secondly because of the nature of his inquiry, which involves *new* uses of concepts and terms and the creation of signs and a notation. A brief summary of this without exposition would convey no meaning. Dr Bion writes: 'Because psycho-analytic theories are compound of observed material and abstraction from it, they have been criticized as unscientific'. He seeks to construct a theory descriptive of 'the practice of psycho-analysis' which uses only 'pure scientific abstractions', a psychoanalytic theory of a kind which does not yet exist. He writes: 'The description I have given is not of a process of thought appropriate to actual contact with the patient. The analytic session is too precious an opportunity for observation to be jeopardized by pre-occupations of the kind my description implies; the theme of this is a scheme to facilitate thinking about analytic work *outside* the sessions themselves.' His work presupposes some knowledge of philosophy and of the philosophy of science in particular. The criticism that he uses a new terminology little related to existing psychoanalytic writing is irrelevant. Whether his particular scheme is satisfactory or not, the attempt has to be made.

Bion constructs his scheme by first seeking to isolate the *elements* of psychoanalysis, i.e. the basic 'functions of the personality', the 'names for sets of actions, physical or mental, governed by or directed to a purpose'. The actual 'element' is an 'unknowable' behind its verbal representation, like Kant's 'thing-in-itself', but knowable by its primary and secondary qualities. In chapter 1 he selects seven such elements, representing each by a sign. The first two are processes: (1) Projective Identification, involving the notion of 'container and contained', hence the sign ♀♂. (2) Reaction between the Paranoid-Schizoid and Depressive positions, the sign being PS ↔ D. In the medium of the operation of these processes, the raw material of experience is worked into forms usable for 'thinking'. The next three elements are the 'passions'. (3) Loving. (4) Hating. (5) Knowing, signs L, H, K, representing 'links between psycho-analytic objects'. (6) Reason, sign R, 'A function intended to serve the passions by leading to their dominance in the world of reality'. (7) Ideas, sign I, to 'represent psycho-analytical objects composed of α-elements'. In *Learning from Experience* Bion made a distinction between

β-elements, the raw material of immediate experience, sensa, and α-elements, or sensa mentally digested and worked up into 'elements capable of storage for use in dream and other thoughts'.

An analysis of 'the reality of psycho-analytic practice' will seek to discover how these elements operate in the statements made by the patient, the Ideas he presents and the feelings associated with them, and in the statements of the analyst. Bion takes a possible statement by a patient, 'I know that you hate me'. These words can express a variety of things, and a scheme is needed to reveal all the possibilities of meaning. He constructs this by asking two questions; (*a*) What is the *developmental level* of the 'thought' the patient expresses? (*b*) To what *use* is the statement put? He suggests seven levels of development of 'Ideas' or 'thoughts' relevant to psychoanalysis, lettered A to G as the vertical axis of a grid, starting with sensa or β-elements (A) and leading to Concepts (F) and a developed Scientific Deductive System (G). He selects six uses relevant to psychoanalysis to which statements embodying these Ideas may be put, numbered 1 to 6 as the horizontal axis of the grid; such as (1) definition, (2) psychic defence, (5) inquiry, (6) production of an effect, a verbal 'acting out' analogous to an action. Thus the grid gives forty-two possible meanings for any statement made in psychoanalysis. To illustrate the two extreme cases, A1 and F or G6: If the analyst thinks the patients' statement 'I know you hate me' is a simple definition (1) of a feeling in the patient, it is a β-element not yet further developed (A), and it is expressed or got rid of by uttering it, 'having the patient no longer burdened by the feeling that the analyst hates him (and) now free to feel that the analyst is his friend'. If, however, the analyst feels the statement represents the hardening of a belief into a fixed Idea (F or G) with which the patient is attacking the analyst (6), then F or G6 means that the patient is convinced the analyst is his enemy, and is about to act on it.

This is a different type of scientific theory of psychoanalysis and should constitute a challenge. I will only state a few comments and questions. (1) Many therapists practice the art of psychotherapy with a minimal interest in theory and most show a marked reluctance to master new ideas. (2) Bion's list of elements calls for critical discussion. It is not clear what he means by an element. It is defined in chapter 1 as 'a basic function of the personality' and seven are listed. In chapter 20, defining 'the elements... as those phenomena whose various aspects fall within the grid categories' he gives a different if overlapping list: Ideas, feelings, association and interpretation, the 'couple' (association *and* interpretation), conflicting pairs (which are not specified) and the axes of the grid. Some may think his views imply an equation of psychoanalysis with Kleinian theory. A fundamental difficulty exists in calling the interaction between the Paranoid-Schizoid and Depressive positions an 'element'. This, it seems, isolates a basic process whereby the psyche develops from unintegrated, split or fragmented states towards unification, integration, wholeness, though under stress the process can be reversed. PS and D illustrate the mixture of observation and abstraction of which Bion complains. They are (*a*) terms in psychopathology for *illness*, (*b*) used by Kleinians to denote 'positions' in *normal* development, and now (*c*) by Bion in a still more abstract way for elemental states of unintegration and unification and processes from one to the other. 'Elements' cannot be represented by terms so loaded with pathological significance. Also, Bion assumes the correctness of Klein's metapsychology that the primary state of the psyche is unintegration, which is disputable. To me, the Gestalt theory that mental development proceeds by differentiation within a primary whole is more convincing, and is a psychological parallel to physical growth. (3) Bion writes: 'The analyst is restricted to interpretations that are an expression of a K relationship with the patient. They must not be expressions of L or H.' Is Bion's a theory purely of psychoanalytic interpretation, or of psychotherapy, or does he assume their identity? If so, it may be thought that this analysis, based so much on Bion's study of disturbed thinking, is too intellectualistic a scheme to be adequate as a theory of psychotherapy. If the analyst may 'know' the patient but not 'feel' anything about him, many will regard that as too impersonal a relationship to be therapeutic. (4) In that case, does not this analysis fail to do justice to the all-important fact that it is 'persons' not 'processes' we seek to study, since it applies to a very imperfectly personal relationship from which 'feeling' is excluded on one side?

H. GUNTRIP

Memories, Dreams, Reflections. By C. G. JUNG. Edited by ANIELA JAFFE. London: Collins and Routledge and Kegan Paul. 1963. Pp. 383. 45*s.*

Contact with Jung. Edited by MICHAEL FORDHAM. London: Tavistock. 1963. Pp. x + 245. 42*s.*

A reviewer, like the writer of a book himself, inevitably brings to his task something of his own point of view—in this case that of one who had occasion to make an historical and critical study of Jung's work at a time when comparatively little biographical material was available. With the publication of Jung's autobiography, *Memories, Dreams, Reflections*, it is natural to ask in what ways this material reinforces impressions reached previously from perusal of his psychological writings, and in what ways is it necessary to revise one's views?

In many ways previous impressions are confirmed. There is, of course, Jung's account of his early and his student years, in traditional European scholastic circles but with roots in the country rather than the town. Of more direct relevance, for an understanding of Jung's work is the increased evidence of his reliance on visual imagery, shown in his reports of unusually vivid dreams. This characteristic seems to have been virtually innate, judging by the descriptions of scenes from the first two or three years of his life with which the book begins. Another dominant tendency in his work—the habit of personification—also seems to have been a basic feature of his own mentality. Even during schooldays, he declares, he had a sense of being two persons. In coming to Jung's account of the years immediately following his break with Freud we find that, while their importance could be estimated previously, the formative influence in the period must now be regarded as springing primarily from a long-continued experience of his own and only secondarily from the experience of his patients. This, probably, is the main difference in emphasis that one would now feel justified in making. At the same time confirmation is given of his interest in the psychotherapy of the psychoses. A still earlier interest, and defection from orthodoxy, in religious matters is reported.

As might be expected, there are many indications of a rather emotionally toned approach to external as well as internal events, e.g. Jung writes in referring to his acquaintance with psychiatry before the reading of Krafft-Ebing's *Textbook.* Many readers, without necessarily wishing to deny the existence of all parapsychological phenomena, would consider this approach to be most apparent in his description of poltergeist experiences. Two interesting omissions are the lack of any detailed reference to his chief at the Burghölzli, Prof. Eugen Bleuler, and to the place in his work and life of family relationships subsequent to marriage.

Contact with Jung is a volume of memorial essays. As Michael Fordham explains in his Editorial Note, papers were invited from members of societies for analytical psychology throughout the world, with the suggestion that contributors might either consider what had been the most significant stimulus they had derived from Jung or give an account of the growth, development and prospects of analytical psychology in their particular group. The mixing of the two types of paper—classification being according to the country of origin rather than the subject-matter—perhaps reduces the reader's sense of unity, and evaluation of the personal tributes may be more difficult if an awareness already exists of the impact which Jung's personality has made upon his followers. The historical essays, including two papers from Germany (K. Bügler and H. Buder) and one by G. Stewart Prince on 'Jung's Psychology in Britain' seem to contain more fresh material, and give rise to the thought that before it is too late to gather records and impressions a further memorial might be made to Jung in the form of a History of Analytical Psychology.

AVIS M. DRY

The Rorschach in Practice. By THEODORA ALCOCK. London: Tavistock. 1963. Pp. xii + 252. 63*s.*

Miss Alcock has at last provided in print the substance of her instruction in the Rorschach method at the Tavistock Institute of Human Relations. As E. L. Trist points out in the foreword, a distinctively British Rorschach tradition has become established over a period of twenty-nine years; Miss Alcock was the introducer, and is the outstanding exponent of the method.

This book, as an account of Rorschach principles and their practical application, incorporates not only developments in the method itself but also recent developments in psychoanalysis and the experimental psychology of perception.

The volume is in three parts. Part I is a thorough account of the method, including administration, significance of scoring symbols, and the assessment of evidence from the interdependent components of psychogram, sequence, and behaviour. Possible pointers towards psychoneuroses and psychosis, as well as intercranial damage, are included. Part II provides detailed case studies, and here Miss Alcock's own distinctive contribution is particularly clear. Part III is an account of a research study of some personality characteristics of asthmatic children.

In her theoretical discussion, the author points to certain basic assumptions: (*a*) The Rorschach is essentially a sampling method, already familiar in science. (*b*) The individual's responses are held to be characteristic of his personality in action in other life situations. (*c*) Perception, as in everyday life, will be influenced by personal 'set' in which unconscious and conscious selection interact. (*d*) The Rorschach is to be regarded for the most part as a test of object relations as defined by Fairbairn and others.

From among the important contributions of Miss Alcock herself, the following are a selection. First there is a sensitive and perceptive discussion of the relationship between the tester and testee which, although bearing on the Rorschach test itself, has wider reference to psychological testing. Next, a very worthwhile addition to content classification is elaborated in the score *Sin*, registering sinister tone. Further, the chapter on 'Conflict and Defence within the Healthy Personality' is of particular importance. It is only too easy to demonstrate pathology in Rorschach responses; the difficulty lies in estimating the nature and degree of personality defence which may counterbalance the disturbances. Here, Miss Alcock's account of certain 'freedoms' of personality is instructive and important. In more general terms, also, her counsel regarding the balancing of strengths and weaknesses in an individual personality, is outstanding. Finally, the importance attached to the individual's perception of the reality of the blot structure is stressed. This concept of blot reality would have merited more extended discussion; it has implications also for other projective tests.

The only slight criticism the present reviewer would make is that in this rich volume there are many important insights rather briefly stated, but to remedy this would of course have meant a longer book. Those insights and their interpretative importance, combined with Miss Alcock's insistence on objective assessment, afford the reader the rare pleasure in one area, at least, of a text-book of psychological psychology. And in conclusion, if the reader should be seeking a really adequate criterion for judging the Rorschach method, then this is the book to read.

FLORENCE MACNEILL

Strukturelemente der Persönlichkeit. By DAVID KADINSKY. Bern and Stuttgart: Huber. 1963. Pp. 106. DM. 18.50.

A union between factor- and psychoanalysis is rare. Mr Kadinsky, therapist and Rorschach worker of long standing, has attempted it. The idea underlying the study is to move away from 'symbolism towards the abstract' and to express psychological events mathematically. Mr Kadinsky extracted five factors from the Rorschach records of 50 people, correlated them with criteria evolved in therapy with 25 patients and interpreted them on analytical lines. He then proceeded to calculate the five factors for each patient's Rorschach record and called the resulting five figures the patient's 'specification equation'. The 25 'equations' were again factor analysed and gave three bipolar factors: father ties; mother ties; castration and compensatory reactions. Mr Kadinsky envisages that it will become possible to calculate at the beginning of therapy what the patient's 'equation' will be at its conclusion.

Mr Kadinsky gives references in footnotes but does not mention the well-known factorial study by Sandler and Ackner in which the authors for good reasons extracted their factors from objectively ascertainable Rorschach scores (content) and correlated the factors with independent ratings of symptoms, previous personality and items of the case history. Mr Kadinsky worked single-handed, is well aware of the subjective element in his research, and wonders himself whether a repetition of his study would produce the same results. He is, however, satisfied that his statistical approach is an independent way of showing essential elements of the personality structure. The reader, whether statistician or clinician, is likely to remain unconvinced.

A. KALDEGG

Emotional Flexibility-Rigidity as a Comprehensive Dimension of Mind. By SIGVARD RUBENOWITZ. (*Acta Psychologica Gothoburgensia*, IV.) Stockholm: Almqvist and Wiksell. 1963. Pp. 256. Sw.kr. 40.

This is an interesting and important study stemming from the group of ideas and hypotheses originally put forward by Jaensch in his book *Der Gegentypus*, and later psychoanalysed and popularized (without much acknowledgement) by the authors of *The Authoritarian Personality*. Rubenowitz has made use of a large and cleverly designed battery of tests (mostly questionnaires), administered to large and varied samples of the Swedish population to ferret out, by means of correlational and factor analytic methods, the underlying dimension of flexibility-rigidity. In addition to such obvious tests as the California F Scale, Rokeach's Dogmatism and Opinionation Scales and the Gough–Sanford Rigidity Scale, he also used an Empathy Test, Child Training Scale, Edward's Personal Preference Schedule, Zulliger's ink-blots, the Rees–Eysenck Body Type Index, intelligence tests, perseveration measures, tests of musical and art appreciation, as well as an autobiographical questionnaire, a study habits inventory and the Strong vocational interest blanks. Examination marks were also available for some student groups as well as ratings of socio-economic class.

It can be imagined that the intercorrelations between these variables, calculated separately for different groups, produce a bewildering plethora of results. In addition, some of the constituent tests such as the F Scale and the battery of perseveration tests were separately factor analysed. The whole operation was carried out efficiently and well and Rubenowitz must be congratulated on guiding the reader through this maze of results without inducing too large a measure of mental indigestion. His central conclusion was that 'in spite of the different character of the two main groups involved in the testing of the main hypothesis, a flexibility-rigidity construct could be established in both of them in accordance with the assumptions expressed in the hypothesis. These findings were sustantially suported by the accumulated evidence from the secondary hypotheses, derived from the nomological network defining the construct and serving to throw light upon it in a more specific way.' This is a fair summary; Rubenowitz has been more successful than his predecessors in defining this very elusive factor. A determined critic, however, might raise three points: (1) The problem of response set has not been handled too well; it might have been better if the major defining variables such as the F Scale and the Dogmatism Scale had been re-worded in such a way as to obviate this factor. (2) It might have been desirable to introduce some measures of a wider framework into the design; there is evidence that rigidity may be nothing more than a particular combination of introversion and neuroticism, and the absence of these variables in the factor analysis may lead to erroneous conclusions. (3) The interpretation of factors is often a little forced. What shall we make of the 'intelligence' factor on page 190 which has a higher loading on a test scored for the number of Ss written in a minute, than on an intelligence test? However, not withstanding these criticisms this is clearly a major contribution and can be recommended to all students interested in social psychology.

H. J. EYSENCK

Measurement in Personality and Cognition. Edited by SAMUEL MESSICK and JOHN ROSS. London and New York: Wiley. 1963. Pp. xi+334. 45*s.*

The main body of the book is a series of papers on the response styles of subjects responding to questionnaires, social situations and cognitive tests. G. G. Stern, whose massive tables throw Robbins's into the shade, makes an attempt to measure the press of a student's environment, that is, the pressures which he thinks the College atmosphere puts upon him. Stern claims to be able to measure press independent of the student's needs and each press is associated with a need of the same name. Within a College and in a sample taken from several Colleges these paired needs and presses have only low correlations but they correlate more highly between Colleges. Edwards, and Johnson and Messick, report evidence that acquiescence comes into play only when social desirability is not operative. Dahlstrom criticizes the circularity of some attempts to prove the importance of social desirability. The discussion of control or ego-mechanisms in perception reveals the concept of perceptual field dependence broadening and rotating in the course of analysis into the concept of general intelligence—a repetition of Spearman's experience. E. E. Jones's paper on the interaction between approval and the need for it stands out from the rest in the use of external criteria and in the experimental manipulation of variables. Cattell contributes a very clear resumé of his research practice and rationale.

J. Ross, in a well-argued paper on levels of measurement, points out that factor analysis may be quite inappropriate for testing hypotheses derived from ego-psychology since the latter postulates mental processes which are better represented by logical operators than by weighted combinations of variables. The presence of ego-control does not determine a person's score in a certain equation, rather does it specify which equation is appropriate to his case.

In many of the papers the reader is oppressed by the height of theoretical skyscrapers reared on nothing more solid than within-group covariance. Holt's criticism of factor analysis from the standpoint of a philosophically sophisticated psychoanalyst is therefore welcome. Tomkins characterizes factor analysts as idealist metaphysicians, to be contrasted with the reality-bound empirical students of the untidy 'given' of humane experience.

The critical papers are generally better than the investigations of response style, which are of limited interest.

KEITH HOPE

Personality and Behaviour. By JESSE E. GORDON. New York and London: The Macmillan Company. 1963. Pp. x + 597. $6.50.

The author is Assistant Professor of Psychology at the University of Michigan and has prepared this text to meet his requirements for introductory courses in personality and human adjustment. He takes personality to be a 'derived phenomenon' bridging several areas of psychology, such as physiological psychology, learning, studies of brain function, psychoanalytic ego psychology and general experimental research on concept formation, and he uses the language of learning psychology to communicate the relevance of the various branches of psychological knowledge to personality.

The book has four sections covering eleven chapters, each of which ends with a 'reprise'. After the Introductory section a section on Basic Processes covers Physiological Backgrounds, Learning, Motives and Thinking. The third section on Contents of Personality deals with the socialization of Primary Drives, of Acquired Drives and of Thinking. The fourth and final section, headed Consequences, includes Conflict, Patterns of Behaviour and a final chapter on Human Nature and Value.

The book covers a lot of material in a fairly superficial manner. The nature of the author's goal in this book makes it inevitable that it should be a 'hotch-potch' which might irritate the teacher but could make some impression on the student. Thus it could be recommended for supervised reading or for use in seminars at about the junior honours psychology level. It is not basic enough to serve as a text-book.

The author has done himself a disservice by editing his text inadequately, with the result that it is repetitious and too long. It is sparsely illustrated, the diagrams are badly explained, and terms are used in a sloppy fashion, e.g. 'Freud tackled the problem of unconsciousness'. For this reviewer the most irritating feature is the juxtaposition of the terms of learning theory and of ego psychology.

R. M. MOWBRAY

Social Behavior: Its Elementary Forms. By GEORGE C. HOMANS. New York: Harcourt, Brace and World; London: Routledge and Kegan Paul. 1961. Pp. xii + 404. 30*s*.

In writing of elementary social behaviour Homans clearly intends the stress to fall on the word 'element': he is concerned with units of analysis at that level of universality where 'there is neither Jew nor Gentile, Greek nor barbarian, but only man' (p. 6). 'I believe that the general features of elementary social behaviour are shared by all mankind' (p. 7). Such notions as roles and institutions he considers already too constrained, specific or culture-bound for his purpose: to provide the conceptual building-bricks which, in different combinations, construct our ideas of vast varieties of social phenomena.

The most elementary terms in which social behaviour is to be described are Person and Other; activity and reinforcement; frequency of interaction and value of interaction. These terms are used in a variety of propositions of which the following are simple examples. 'The more often within a given period of time a man's activity rewards the activity of another, the more often the other will emit the activity.' 'The more often a man has in the recent past received a rewarding activity from another, the less valuable any further unit of that activity becomes to him' (pp. 54–55). The difficulties arising from such propositions are fairly acknowledged.

What must be the underlying aim of the book impresses itself on the reader in the first hundred pages. It is of course, partly, to present as attractively as possible a set of concepts and propositions for systematic application to apparently disparate social phenomena. But beyond this we glimpse the old hope of a Promised Land of universal social science. Homans is laying down foundations for a formalization of the study of interpersonal relations, by some combination of behavioural psychology (Skinner's pigeons providing him with his original paradigms) and mathematical economics. Since H. A. Simon made a limited attempt of this kind in an article reproduced in his *Models of Man*, a whole cohort of Joshuas has had time to fit itself for the purpose. Homans himself in the present book attempts a formalization of the relation between a man's satisfaction with a reward and the amount of behaviour he will engage in to obtain that reward.

Homans begins with the relatively concrete, simple and, in the literature, traditional notions already referred to, but he soon develops his theme in a novel and stimulating way. Social approval is treated as a 'generalized reinforcer'. Social influence is considered in terms of the costs and rewards which accompany similarity of opinion. Conformity, co-operation and competition are discussed in commensurate terms. It is perhaps difficult to convey to those in other fields the pleasure of seeing these topics all discussed in terms of the same elements.

More general still, more exciting but more controversial also, is the discussion of justice, which, in terms of elementary social behaviour, may be seen as a certain proportionality of rewards, costs and 'investments'. 'As a practical matter, distributive justice is realised when each of the various features of his investments and his activities, put into rank-order in a comparison with those of other men, fall in the same place in all the different rank-orders' (p. 264). Obviously this is a useful point of departure for the study of small groups and their sometimes rather forced hot-house emotions, but one feels here, as in the later discussions on authority and equality, the conservative undertow so often at work in social theory. Such matters as social justice, and changes in the distribution of rewards or in the criteria for reward, are dealt with by omission. 'Esteem', 'exchange', 'satisfaction', appear in the index, but not 'exploitation', 'politics', or 'power'.

But it is wasteful to expatiate on ideological points in this kind of work. Science is a co-operative enterprise in which deficiencies here can be made good there. *Social Behavior* makes sense both in its own right and as providing a common language for studies, both empirical and theoretical, written in terms of less elementary units of behaviour. Seriously considered, say in a joint seminar, the book could prove a turning point in the collaboration of economists with psychologists and other social scientists.

JOSEPHINE KLEIN

Social Pressures in Informal Groups. By L. FESTINGER, S. SCHACHTER and K. BACK. London: Tavistock. 1963. Pp. x+197. 38*s*.

The history of social psychology is short, and a good book can become a 'classic' a few years after its appearance. 'Social pressures...' deserves the status. First published in 1950, it is now reissued with a short introduction by T. M. Newcomb.

The book describes an investigation of a housing project for ex-service students at M.I.T. The best known and least equivocal finding was that the formation of social groups (with all that this is shown to imply for action) is strongly influenced by the spatial layout. The *fact* of this is obvious, but its *degree* was remarkable. In such a homogeneous community, friendships are appreciably affected by differences in the separation of houses measured in feet and by the directions in which the houses in a court are faced. This 'micro-ecology' has inspired other work in housing, discussion groups, offices, classrooms, social clubs, children's holiday camps and student dormitories.

However, it is not through spatial studies alone that the book has become a leading reference, and indeed, only one of its nine chapters is wholly devoted to ecology. The remainder describes the application of a variety of techniques to discover how the processes of group pressure operated to determine attitudes and behaviour towards a *tenants' organization*. The concepts used—group structure, cohesiveness, power, communication, deviation, etc.—were already the currency of group dynamics and there is no particular innovation that could alone explain the impact of the book. Looking back, there are perhaps three important factors. First, it is not only a lucid description of a field situation, but is translated into testable, refutable propositions, drawn together into a final theoretical summary. Secondly, it was presented by students of Kurt Lewin soon after the tragic but critical milestone of his death, and thirdly, the authors have since built upon its foundation some of the best small-group studies of the past decade.

TERENCE LEE

Behaviour in Public Places—Notes on the Social Organisation of Gatherings. By ERVING GOFFMANN. Glencoe, Ill.: Free Press; London: Collier-Macmillan. 1963. Pp. vii + 248. 45*s*.

Erving Goffman is an American sociologist who has been developing a brilliant and original line of study over the last few years. This book draws together and re-presents his main contribution. It rests partly upon foundations of observation of social behaviour which he made in the Shetland Islands in 1949–51 and in psychotic wards in St Elizabeth Hospital, Washington, in 1955–56. He draws further material from casual observations and from etiquette books. His subject matter is the ritual or conventional regulation of engagements or encounters between people in social gatherings (i.e. sets of people in each other's immediate presence), in situations (i.e. people in a place), on occasions (for weddings, parties, work). His predominant and fundamental case is that it is possible and desirable to consider situations and their rules as something distinct from studies of individual personality, groups and roles. This case is well argued and supported by the actual conceptual scheme he is able to present. The scheme rests on a postulated universal need to present to others a consistent 'self' or face, and to be assured that others are likewise concerned. He deals with the means by which a person can indicate to others his involvement in a situation, guard himself from involvement, offer himself as accessible to engagement and protect himself from being engaged unwillingly. He deals with these in situations ranging from the casual passing of strangers to the limitations which may be placed on two or more people devoting their attention to private affairs in public.

Goffmann shows how the rules or 'situational proprieties' can be seen as defensive of the gathering itself and that gatherings thus regulated form a web of trust that others will not endanger a person's physical and social inviolability. He comments that the inherent dangers are not commonly realized, but that it is only in gatherings that people can be violently or sexually assaulted and for instance be accosted by requests to talk or be seen by others to be on close terms. Thus, he argues that however trivial the rules and the etiquette may seem, their social and psychological value is indicated by the deep embarrassment which breaches can occasion.

A special concern of the author is to apply his scheme to the field of mental illness. He argues that the behaviour of psychotic patients can be subject to conventional regulation and can express something about their relations with their situation, either respect for it or rejection of it. This is fairly straightforward and his scheme offers increased precision in understanding the communications patients are making. Two further main points are at different levels of abstraction and are not consistently separated. One is about the reaction of the public and of psychiatrists to psychotics in terms of his scheme. He wishes to show that the diagnostic and treatment procedures to which psychotics are subjected are in part protective reactions against the offences to situational proprieties committed by psychotics. Their behaviour is designated as symptomatic of illness and some of their behaviour follows from this. The other point is about the relations between his sociological scheme and psychological theory. He argues that none of the behaviour of at least some patients need be considered symptomatic of a disturbed mind, and that no recourse to psychological theory is called for.

There are two fashionable views of psychosis. One is that it is directly symptomatic in the way that abdominal pain is symptomatic of appendicitis and the other is that it all consists of odd behaviour only designated as psychotic because it offends social rules—which depends on the rules, not the individual. Goffmann tends to assert the second as an antidote to the first, without quite committing himself. He tends, in Piagettian terms, to assert that behaviour is not simply accommodation, but that it is all assimilation. He is unable or unwilling to maintain this view when faced with organic psychosis. The antithesis is false and his arguments on this count leave much to be desired.

By focusing attention on gatherings rather than roles or groups, he brings us close to a scheme linking an analysis of the normative framework of behaviour and the needs of people engaging in it. He frequently postulates particular needs, but it is not his concern to adopt any systematic account. Nevertheless, his account is of particular interest because of its potentially close contact with any psychological system which offers an account of the self, of anxiety and defence and of object relations. He offers a conceptual scheme which brings order into a field of observation which has been of psychological and psychiatric interest for a long time. G. HUTTON

Determinants of Infant Behaviour. Vol. 2. Edited by B. M. Foss. London: Methuen. 1963. Pp. xii + 248. 45*s*.

This record of the Proceedings of the Second Tavistock Seminar on Mother–Infant Interaction shows a shift in interest away from studies of neonatal behaviour, institutional care and other forms of maternal deprivation, which characterized the first Seminar, towards more actual observations on mother–infant interaction and theoretical discussions on the development of attachments in humans.

The animal studies have the usual advantages over the human, in that the infants grow up more quickly, and can be subjected to more rigorous controls. Harlow's 'unmothered' monkeys (raised in isolated cages or on cloth surrogates), whose infant behaviour was described in Vol. 1, are now seen as adults, with disturbances of mating behaviour (if they have also lacked infant companionship), and in maternal attitudes. Their infants, although ignored or roughly rejected, nevertheless show a strong urge to cling to their mothers, quite apart from the desire for food. This seems to support Bowlby's theory of instinctual responses in attachment development. Among further studies relevant to the 'Maternal Affectional System', Harlow's observations on mother–infant interaction in a controlled playpen situation provide evidence in favour of 'monotropy' of infant attachments. This is borne out by Rowell's observations on monkeys in more natural 'family' groups.

Ainsworth's delightful account of 'The Development of Infant–Mother Interaction among the Ganda' describes twelve different patterns of attachment behaviour appearing in orderly fashion at different ages, thus favouring Bowlby's instinct theories once again. She also supports the concept of innate 'monotropy'. The main opponent of this idea is Schaffer. His research uses a much narrower criterion of attachment, which perhaps explains the discrepancy. Ainsworth finds the chief determinants of attachment in the total amount of care and interest given by the mother, rather than in any specific method of child-rearing, and also stresses the importance of the infant's own activity in forming relationships.

This latter aspect forms one of the chief topics of the remaining papers. Prechtl, for example, reports that mothers of children with signs of minimal brain damage at birth are often upset by their infants' abnormal responses (such as hyperexcitability), and tend to develop attitudes of anxiety and over-protection towards them. Wolff, in the first of two papers, makes a very intensive study of smiling, one of the most obviously social responses, in eight infants, and in his second article describes the family relationships of one of them in greater detail.

Rheingold's short contribution on a method for 'Controlling the Infant's Exploratory Behaviour' indicates that changes in visual and auditory stimulation are sufficient reinforcement for a manual response to a ball, and suggests a similar explanation for the reinforcement of smiling to human stimuli. This theory seems to be accepted by Schaffer in his discussion of 'Attachment Behaviour', where he outlines three developmental stages, relating these to Piaget's ideas on the attainment of reality concepts.

Ambrose, in 'The Concept of a Critical Period for the Development of Social Responsiveness', gives quite plausible grounds for entertaining such a hypothesis, based on a comparison between smiling in humans and 'following' in birds, both of which serve to keep the child in closer contact with the mother, and perhaps thereby reduce anxiety. (Could a further parallel with clinging in monkeys be made?) His own evidence from observations on the pattern of development of smiling is, however, by no means conclusive, although ingenious and suggestive. As Ambrose himself indicates, the type of experiment which might be more convincing, such as Harlow has undertaken with monkeys, where opportunity for social interaction is totally lacking for a certain period, is neither possible nor indeed desirable in human beings.

This brief outline of some of the points brought out in this stimulating collection of papers cannot do justice to the quality and variety of research undertaken, and the book can be recommended to the serious student of child development. The theoretical approach varies, some authors seeming to support Bowlby's concept of instinctual responses, others importing further ideas from ethology, while others again confine themselves more closely to learning theory. In all cases the infant is conceived as having a need for stimulation and activity beyond the mere satisfaction of his organic needs.

A. E. NAUGHTON

Child Psychology. The 62nd Year Book of the National Society for the Study of Education. Part I. Edited by H. W. STEVENSON. London: University of Chicago Press. 1963. Pp. x+550+vi. 48*s.*

This book is designed, not as a text-book or manual covering the whole field of child psychology, but as a critical review of research findings and thought in various areas of the subject, chiefly over the past decade. The contributors, not all of whom may be familiar to British readers, are clearly all competent in their own fields, and their combined endeavours make an authoritative and highly valuable source book.

The first two chapters deal respectively with the biological and sociological correlates of child behaviour; then follow chapters on language development, perceptual development, learning and thinking in children. In place of the usual rubrics of social and emotional development we find chapters on moral development and identification, dependence and independence, aggression, achievement and anxiety. In an admirable concluding chapter Bronfenbrenner surveys the changes in child psychology over the past thirty years, comparing this volume systematically with the handbooks of Murchison (1931) and Carmichael (1946) and offering stimulating comments on the changing assumptions, concepts and methods adopted by adherents of learning theory, field theory, psychoanalysis and Piaget's thought as each concerns children. While there are some signs of convergence, we are still far from a science that can embrace all aspects of the child's behaviour and experience. The editor in his short introduction suggests that longitudinal studies, observational methods, and a developmental orientation have been largely replaced by short-term experimental studies of the effects of particular variables on child behaviour. It is by no means clear, however, that the need for the former types of study is past; indeed they offer the only counterpoise to the fragmentation inherent in the latter approach. Perhaps indeed the form imposed on this book has resulted in the neglect of those more global studies of child development which still flourish, and also of several recently launched observational studies focused on the details of interaction between persons (mother and child, teacher and child, children in groups) in natural settings.

It also forces the critic to ask what else has been left out. One accepts the decision to omit the whole area of deviant behaviour, and also the practical applications of research findings (many of which invite replicative experiments in practical settings by any interested psychologist). Much more questionable is the tacit neglect of the whole imaginative life of childhood. The index contains one reference to projection techniques, none at all to play, art, fantasy or imagination. This does seem an extraordinary deficiency in a major work of child psychology, in view of the growing recognition of the importance of these processes in healthy personality development.

Although the lists of references following each chapter are valuable, they contain no indication of the page on which each reference is discussed, nor is there any author index, so that it is quite difficult to find where a particular author's work is mentioned. Despite these lacks, the scope of the work reviewed is such that this will soon become an essential reference book.

TERENCE MOORE

Language, Thought and Personality in Infancy and Childhood. By M. M. LEWIS. London: Harrap. 1963. Pp. 256. 18*s.*

This book is a comprehensive account of the language development of children, relating it closely to intellectual, social and emotional growth. The first part of the book deals in considerable detail with the beginnings of speech in early infancy. The sequence of the earliest sounds, their relation to bodily comfort and discomfort, the sound patterns evident in babble language, the growth of meaning and of perception and conception in relation to language are vividly described. Great importance is attached to the child's contact with his mother, her stimulation through everyday conversation with him, the approval and disapproval shown by her at his first attempts at vocalization. This is contrasted with the limitations imposed on the deaf child and the child brought up in an institution, who show clear intellectual and linguistic retardation. The quotations of infant speech, many of them from the author's own records, serve to illustrate the argument of the book most effectively. Information concerning the intelligence as well as the age of the children quoted would have been valuable.

The second part of the book continues the survey of language development during early childhood, drawing especially on Susan Isaacs's records from the Malting House School. The discussion of the role of language in the child's growth of reasoning, the vital part of play and social intercourse in increasing vocabulary and verbal fluency, the way in which real problems incite children to co-operate in solving them and so to carry forward conceptual thinking—all this is extremely interesting. Ethical development is described in some detail. This might more appropriately have been termed social and emotional development in my view, but the author wished to stress particularly the influence of the adult attitude of approval and disapproval on language development.

The third part of the book describes the gradual growth of abstract and logical thought in relation to language and draws conclusions from the experimental work of Piaget and others which further elucidates his argument. But perhaps the most fascinating section of this last part is the account of the social life of later childhood, the verbal play, the group language, the rituals of children's games and the secret rhymes aimed at challenging adult authority and an indication of the child's 'need to lay siege to the barriers that exclude him from the adult way of life'. Many of the illustrations are drawn from the work of I. and P. Opie.

The whole is an up-to-date survey of the most important work on language development during infancy and childhood, emphasizing the links between the child's past and present experience and stressing the immense importance of social intercourse, with parents, other adults and especially with contemporaries. It is written in clear, mainly non-technical language and is a worthy addition to the list of other notable books by Prof. Lewis.

AGATHA H. BOWLEY

A Study of Children's Thinking. By MARGARET DONALDSON. London: Tavistock. 1963. Pp. viii + 263. 35*s*.

This book is a clinical study oriented primarily towards the improvement of intelligence test items. The first two chapters deal with the prediction of ability and with related methodological problems. The third chapter is an excellent discussion of the methods used in studying thinking and the externalization of thought processes. The merits and limitations of Piaget's approach are discussed, as well as the psychological implications of the views of Peirce and Ryle. This chapter also contains a balanced account of the pros and cons of the 'thinking aloud' method. It is curious, however, that the work of Herbert Simon and his associates, who have used this method with such success, is not mentioned here.

The next five chapters are the meat of the book: a detailed investigation, conducted on Piagetian lines, of the kinds of error which occur when children of different ages attempt to solve different kinds of logical problems. And here the experimentalist may be distracted by the very free use which is made of illustrative material in the form of protocols. Much of this material might well have been classified in appendices and more space devoted in the text to a concise exposition of the points which the protocols are intended to illustrate.

The final chapter summarizes the research and discusses its theoretical implications. A postulate set is presented which relates the three main types of error, i.e. structural, arbitrary and executive, to intellectual development. (This chapter, however, cannot be understood without reference to the previous ones because of the symbolic terminology which is maintained throughout the book.)

The author claims only to provide good hypotheses rather than to test them. The book is not easy reading, but it attempts to do something which has not been done before: to look at children's reasoning in relation to diverse theoretical positions. The relevance of the work of Piaget, Goldstein and Scheerer, and Shannon and Weaver is discussed in a way which is illuminating and thought-provoking. The author's most original concept, however, is the distinction which is drawn between three types of error in reasoning, and the hypotheses based on this distinction deserve to be tested by experimental techniques.

In conclusion, it may be said that this research has practical implications for the constructor of intelligence tests, but it should also be regarded as important source material for a systematic account of the development of deductive reasoning.

P. C. WASON

Speech and Thought in Severe Subnormality. By N. O'CONNOR and BEATE HERMELIN. London: Pergamon. 1963. Pp. xi + 122. 25*s*.

In 1957 this reviewer had the task of surveying experimental work on learning in imbeciles; he was able to discuss about fifteen good studies. The present book ends with over nine pages of references, a considerable proportion of which concern recent experimental investigations of low-grade defectives. There has, then, been something of an explosion of interest in this field during the last few years. While applied psychologists in America continue predominantly to explore methods of educating the mildly subnormal, a number of workers from the field of comparative psychology have seen in the severely retarded human an important subject for research. Meanwhile in Russia A. R. Luria (who provides a foreword) and his colleagues, investigating the role of language in the regulation of behaviour, have also devoted considerable effort to studying children displaying abnormalities of function or neuropathological conditions. The Russian school has obviously been the source of much inspiration for the writers of this book, who, however, bring to the problem a more sophisticated method of analysis than is often found in the writings of Soviet psychologists.

In a notable first chapter, O'Connor and Hermelin consider trends and developments in the field of general psychology relevant to a detailed investigation of the thought processes. Here they make the charge that psychometrics and I.Q. studies have retarded the investigation of strategies and mechanisms in thinking. There is little doubt that this is so, but unlike many similar critics, these writers do not throw the baby out with the bathwater by dismissing the psychometric approach.

The second chapter on learning and problem solving is also characterized by a careful historical analysis which indeed permeates the whole book. Curiously enough, however, the authors abandon this method when they reach the research work of the last few years with which they are mainly concerned. Why is it, for example, that Soviet work has always stressed the disabilities of the imbecile, and the rift between primary and secondary signalling systems in these subjects, generating even greater gloom about them than seems justified? Why is it, on the other hand, that British researchers have been rather more concerned with identifying assets than deficits in imbecile learning? The striking dissimilarities between these findings call for much closer inspection than is given here. And even in the present state of knowledge some explanatory hypotheses are possible, the duration of the learning stimulus being one of them.

Subsequent chapters are concerned with visual perception; words and communications; thought and language; cross modal coding; recall and recognition; mongols and other imbeciles; and finally a discussion of theoretical and practical implications. The authors show, for example, that some learning difficulties can at least partly be overcome by the manner in which the task is presented. Part of the defective's lack of capacity to handle symbols comes from a reluctance to use them. When forced to do so, performance improves. The dissociation between speech and action postulated by Luria may well be too pessimistic a view. The authors also point to the relative efficiency in transfer and retention of learning in their subjects, while underlining deficits in expectancies and attention.

Apart from the stress throughout upon experimental method and the avoidance of premature, all-embracing theory, the excellence of the book lies quite largely in its emphasis on processes of cognition. Nor are these processes regarded as necessarily immutable; the conditions for development and change in severely handicapped persons are generally explored. In criticism it would be fair to make two points; the first unfortunately applies to the majority of learning studies, be they animal or human. The authors only carried out short, discrete experiments and give few details of whether their subjects were ever used for more than one such study—a vitally relevant factor. In the reviewer's opinion, this book prepares the way for what might prove to be even more rewarding, longer-term investigations of similar individuals. This should lead to greater knowledge of both the mechanisms of development and the experimental modification of the processes here discussed. A further point is the authors' failure to defend their somewhat controversial policy of using as normal controls school-children whose intellectual and socio-economic position is unstated, and whose mental ages were inferred from their chronological ages.

O'Connor and Hermelin's book is the first ever produced which systematically considers the psychology of the severely subnormal. As such it will prove to be a landmark in the history of mental deficiency research, and furthermore, its relevance to general psychology is obvious.

A. D. B. CLARKE

Experimental Psychology. Edited by A. SUMMERFIELD. *British Medical Bulletin.* Vol. 20, No. 1. London: The British Council, Medical Department. 1964. Pp. 86. 30*s.*

Fifteen psychologists, who are actively engaged in experimental work, have each written a paper dealing with some aspect of their current researches. The topics considered include learning, attention, perception, skill, sensory deprivation, neurology and behaviour, sleep, ageing, and drugs. Although the symposium does not pretend to give an exhaustive coverage of experimental work in psychology, it is welcome on two counts. First, it presents the reader with a lively sample of the variety and high quality of research being carried out in this country. Secondly, as befits the nature of this symposium, the papers illustrate a broadly biological approach to the study of psychological functioning. This approach can best be indicated by quoting from the editor's introduction. 'Papers...relate behaviour to ideas of communication and control that stem from engineering. Both men and animals function in important ways as systems for handling information; derived from sensory stimulation, information is transmitted and coded so as to control decisions and actions. The biological parallel is with the central nervous system. These ideas serve therefore the co-ordinate knowledge about behaviour in ways harmonious with developing knowledge of neurophysiology.' In short, the symposium serves not only the short-term function of keeping the reader up to date about developments in several areas of research, but also serves the longer-term function of exemplifying an ethos which promises to share a flexible yet unified attitude to the subject matter of psychology.

I. M. L. HUNTER

Vigilance, A Symposium. Edited by D. N. BUCKNER and J. J. MCGRATH. London: McGraw-Hill. 1963. Pp. 269. 69*s.* 6*d.*

This book is an example of the valuable contributions which symposia can make by bringing people together, provided that the numbers are kept small, and the participants allowed ample scope. Only eight major contributors are listed, although some of the papers are given joint authorship. Each of the major authors is allowed both a paper presenting recent research, and another to review his own theoretical position at present, and how he sees it fitting into those of the other participants. The result is a real summary of most of the important recent work on the somewhat mysterious problem of why people show decrement in performance when watchkeeping—or in some cases whether it is really true that they do. The symposium is invaluable as a source book. And for once the discussions between the participants have been so well edited that the reader can get a feeling of things being thrashed out, and an indication of which current approaches seem most promising. (It is interesting that the only avowed 'anti-theorist' of the operant school seems both in the presentation of research and in the discussion to be far more rigidly bound by his orientation than the more traditional 'model makers'.)

The names of the major authors—Bakan, Baker, Broadbent, Buckner, Holland, Howland, Jerison, and McGrath—give an idea of the importance of this collection of papers. The book is in the well-known McGraw-Hill series on psychology, and maintains the high standard of production which we have come to expect from the series. For anyone with interests in fields that impinge on the problems of vigilance, arousal, signal detection and such problems, this is an essential text, and one which will both save a great deal of literature searching and give a great deal of stimulation.

NEVILLE MORAY

Experiments in Mental Suggestion. By L. L. VASILIEV. Church Crookham, Hants: I.S.M.I. 1963. Pp. 179. 30*s*.

This is a remarkable work by the Professor of Physiology in the University of Leningrad. The preface reviews the progress of parapsychology in Russia since 1922. At first there was widespread agreement that the many positive results obtained must be due to brain 'radio' waves—i.e. that telepathy had a respectable physical basis. However, three reports published in the years 1934–37 showed that the most careful screening by metal of the agent had no effect on these positive results, thus throwing the electromagnetic theory of telepathy into serious doubt. The war disrupted this work and it only started up again quite recently. In 1960 the University of Leningrad set up a laboratory for the study of telepathic phenomena and experimental work is in progress not only in Russia but in other 'People's Democracies'.

The book consists mainly of full descriptions of Prof. Vasiliev's pre-war work (with contemporary comment), centring on the mental suggestion of motor acts in hypnotized subjects, on transmission of visual images and on telepathic induction of sleeping and waking. However, in this early work the controls set up were not always adequate. In suggesting motor acts the agent and percipient were in the same room; in the experiments on transmission of mental images they were in different rooms but no control seems to have been made of the judgement of whether there was a match between agent's and percipient's pictures. The experiments using sleep were better controlled, with agent and percipient in different rooms, with and without wire mesh screening. The best results were obtained with double metallic screening although statistically significant results were obtained in all experimental situations. Positive results were also obtained with up to 1700 Km. separating agent and percipient.

The last chapter contains a review of contemporary work in parapsychology. It is indeed remarkable that a distinguished Soviet scientist should have to haul G. Price over the coals for his article in *Science* (**122**, 389, 1955), wherein he admitted that the evidence for telepathy was now watertight—but, he concluded, as telepathy was scientifically impossible, all parapsychologists were therefore frauds, liars or engaged in some vast world-wide conspiracy.

Vasiliev concludes with an interesting cybernetic analysis of telepathy. He discusses various theories of the electromagnetic basis of telepathy but cannot accept any—in the end he quotes Pascual Jordan's suggestion that telepathy may depend on something more akin to gravitational forces. The translators are to be congratulated on a most useful and well executed piece of work.

J. R. SMYTHIES

Experimental Psychical Research. By ROBERT H. THOULESS. Harmondsworth, Middlesex: Penguin Books. 1963. Pp. 148. 3*s*. 6*d*.

Dr Thouless is among the distinguished few who have made important original contributions both to experimental psychology and to experimental parapsychology and the present volume is written with so much mature good sense that I can think of no better introduction to the topic that I would want to recommend to those of my colleagues who may still be *psi*-shy. It is true the book does not quite seem to have made up its mind whether it is trying to be a handbook for would-be experimenters or just another brief survey of the field, but if it is approached as primarily a handbook with historical object-lessons the reader will not go far wrong.

The chapter on Marion (the late stage-telepathist) should be of special interest to psychologists because it demonstrates the difficulties in trying to discriminate between the hypotheses of ESP and hyperaesthesia. I had assumed that S. G. Soal had proved conclusively in 1934 that Marion's tricks depended exclusively on his unconscious sensitivity to subliminal sensory cues, but apparently Thouless and Wiesner did some further work with him ten years later which reopened the question, although unfortunately it was never completed or published.

While Dr Thouless does not minimize the formidable obstacles that lie in wait for those who are genuinely concerned to get at the truth in this baffling and often exasperating science, his book contrives to suggest that the effort is worth while.

J. BELOFF

The Science of Animal Behaviour. By P. L. BROADHURST. Harmondsworth, Middlesex: Penguin Books. 1963. Pp. 143. 3*s*. 6*d*.

Dr Broadhurst deserves to be congratulated for a delightfully written, yet provocative, addition to the Pelican Psychology Series. Although this is a short book, it sets out to describe the principal methods of investigating animal behaviour and to explain the rationale underlying comparative psychological studies. As an introduction, the book succeeds admirably on both counts and may be recommended as useful reading material for the first-year student.

The first chapter, 'Why study animal behaviour?' illustrates the nature of the relationship between studies of animal behaviour and human experimental psychology; an excellent discussion of the moral and legal responsibilities of those who work with animals is also given. To some extent, however, this chapter reflects an emphasis upon purely practical reasons for choosing animals rather than human subjects for behavioural study and it would be a pity if this created the impression that all such experimentation is merely a convenient extension of work with human subjects.

The next two chapters provide a good introductory account of methods of investigating animal behaviour in the field and in the laboratory. Many of the elegant techniques now available (for example, Blough's method of determining dark-adaptation thresholds in pigeons) are neatly described. The fourth and fifth chapters are concerned with innate and acquired components of behaviour, respectively. Much attention is given to reproductive and parental behaviour in the first of these, while the second considers, from a methodological standpoint, classical and operant conditioning as well as the relationship between motivation and performance. Chapter 6 deals with abnormal behaviour, including a description of Pavlov's original work, studies of avoidance behaviour, psychosomatic disorders (Brady's 'executive' monkeys), and addiction to drugs. An impressive feature of the book as a whole is the way in which the author draws upon diverse experimental findings to illustrate a methodological problem or some particular facet of a behavioural process.

The last section is thoroughly amusing and provocative. Effectively cued by Skinner, the author considers some of the possible applications of animal behaviour studies to human society. We learn, reasonably enough, that pigeons might be employed to do postal sorting or as production-line inspectors. The suggestion that chimpanzees might be trained as engine-drivers leads one to the conclusion that Dr Beeching could look elsewhere for his economies!

ROGER STRETCH

Handbook of Mathematical Psychology. Vol. 1, Chapters 1–8. Pp. xiii + 491. 80*s*.
Readings in Mathematical Psychology. Vol. 1, Pp. ix + 535. 68*s*.
Both books edited by R. D. LUCE, R. R. BUSH and E. GALANTER. New York and London: Wiley. 1963.

The first of these books presents a series of traditional psychological topics treated from a mathematical point of view, though some of the chapters are primarily methodological. The second is designed as a companion volume of source materials and consists of selected papers describing major contributions to the progress of mathematical psychology in the last fifteen years. An astonishing amount of work is covered in these two volumes, especially in view of the authors' claim, which is entirely justified, that less than a quarter of the material covered in the three Handbooks (Vol. 2 and Vol. 3 are to appear later) could have been written in 1950).

The number of people in this country who could read these two volumes with profit is small. Most of those with a good understanding of the necessary mathematics have either little knowledge or little interest in the psychological topics covered, and most of those psychologists who are interested in the material have insufficient mathematics. Not that the necessary mathematics is all that difficult, but it is cast in a modern vein and might appear incomprehensible unless the reader is acquainted with certain aspects of modern mathematics (set theory for example).

Chapter 1, on Basic Measurement Theory; Chapter 7 on Computers in Psychology; and Chapter 8 on Estimation and Evaluation, in the *Handbook* are particularly interesting. This latter, written by Dr Bush, assumes that the reader is well-versed in statistics and goes on to argue for a 'common-sense' approach to the use of statistical method. It summarizes those parts of statistics most relevant to mathematical model testing and suggests that the average psychologist is far more familiar with hypothesis testing than with parameter estimation, although the latter is equally important.

Few individuals can be expected to buy these relatively expensive books but they should be on the shelves of the psychological section of every library. They are indispensable reference books in a field of work which is growing rapidly in the United States, and which may prove of great importance in the immediate future.

JACK WRIGLEY

The Eye. Edited by HUGH DAVSON. New York and London: Academic Press. 1962. Vol. 2: *The Visual Process.* Pp. xviii + 796. 157*s.* Vol. 4: *Visual Optics and the Optical Space Sense.* Pp. xviii + 432. 100*s.*

This work can be thought of as an advanced text-book in the field of vision, where it maintains a high standard of technical information and is sufficiently exhaustive to be regarded as a 'handbook' to guide and stimulate the interest of those who wish to embark upon research in vision. Of the four volumes in this treatise the two noted above are of special interest to psychologists.

The publishers claim that no previous knowledge is required to follow the arguments in the various subjects. This claim is not exaggerated, as each particular field is introduced by a chapter or chapters dealing with fundamental ideas, and giving definitions of the concepts employed. After this initial gentle introduction the pace quickens and the enthusiasm of each writer for his particular field sustains one through the more difficult passages. Enthusiasm here and there is tempered by some sobering remarks that evidence obtained in a laboratory under strictly experimental conditions may be very difficult to correlate with visual performance in general.

Vol. 2 is the largest of the four books; and many of the chapters are written by internationally known specialists. We have Dartnall's account of the photo-biology of the visual process and Granit's excellent description of the complexity of the Visual Pathway.

There are also fine expositions of the many aspects of the visual function in man by Marriott and Pirenne covering such topics as the Nature of Cones and Rods, Dark Adaptation, Visual Acuity, Colour Vision, etc. This volume is well documented, with extensive references after each specific chapter. It is a pity there is no mention of the fascinating topic of acquired dyschromatopsias in this volume, especially so as in this publication emphasis is laid on its being up-to-date; and also because this field of study opens promising avenues to our understanding of the visual process. It might also have been better to omit or revise some of the more speculative work, such as that described in the chapter on 'The Two-Colour Threshold Techniques of Stiles'.

Vol. 4 essentially deals with two related subjects, viz. 'Visual Optics' and the 'Optical Space Sense'. The first part of the volume is an excellent introduction to this rather difficult subject-matter, the individual chapters being clearly written.

The second part of Vol. 4 can be recommended to all serious students of psychology. This part is well expounded by Ogle, and should be a help to those who find that the psychologist's approach to this subject-matter leaves many problems either ignored or unanswered. Like its predecessor, this volume has well classified references, gives useful hints on possible future research and indicates apparatus suitable for the more advanced study of some phenomena in perception.

R. LAKOWSKI

Pain. Its Meaning and Significance. By F. SAUERBRUCH and H. WENKE. London: G. Allen and Unwin. 1963. Pp. 151. 25*s.*

This book consists of six sections dealing with the physiology, psychology, philosophy and treatment of pain. The treatment of each subject is somewhat elementary, and the medical part would form interesting reading for those medical students who like reading around their subject —for what is said, though simple, is often important. The psychological and philosophical portions tend to lapse at times into existentialist worries about the 'meaning' of pain, but much that is said is shrewd and shows compassionate insight into human suffering—a quality that does not necessarily thrive in this biochemical age. The sections on 'emotional pain' are of particular interest. Less palatable to Anglo-Saxon readers is the Nietzschean conclusion 'Pain, need and suffering can lead through distressing experiences to greatness and new hope, and to joy in the future. What physician has never witnessed a change in personality after great suffering, or never seen new vitality well up precisely because pain and suffering have burned away the dross?' It seems very difficult to say whether the constant exercise of will power necessary to fight great pain—physical or mental—really does so strengthen the will. Perhaps this could form an interesting field for research. The translation (by Edward Fitzgerald) is lucid.

J. R. SMYTHIES

Psychological Depression. A Common Disorder of Personality. By MURDO MACKENZIE. London: Churchill. 1963. Pp. vii+208. 21*s*.

The medical psychotherapist is a much lonelier worker than any other medical man. In dealing with his patients he has to rely on himself alone. He has either to borrow his conceptual tools from others by becoming a member of a school, or he has to evolve his own concepts. The author of this book has done the latter, although he has not been uninfluenced by current psychodynamic theories. He claims that contrast and conflict between the components of personality, whether individual or group, underlie and accentuate mental conflict within the individual. These components he calls Immediacy and Deliberation in the individual, and Simplification and Amplification in the group. The psychotherapy based on these concepts is called contrast therapy, which the author described in an earlier book, *Contrast Psychology*. His new work deals with the common syndrome of persistent anxiety and depression. It is addressed to readers professionally engaged in helping sufferers from depression, but also to the patients themselves. The term psychological depression denotes a clinical syndrome characterized by 'reaction to persisting anxiety by personality denigration'. It requires to be treated by psychotherapy. It has to be distinguished from apathy or inertia, which respond to physical treatment.

The author has found Galton's classification of personality according to imagery useful. Contrasts between those personality types may also contribute to anxiety and depression. The patient has to be made aware of these unconscious conflicts. It is essential to take note not only of the individual but also of the community personality.

The author applies his personality psychology to historical personages and to national mentalities. He illustrates with the help of case histories and dialogues how his treatment works and how he helps his patients to 're-direct and synthesize the psychic forces at their disposition into a simple, creative, positive whole'. Like the early Freudians, he attributes his successes to insight into mental processes rather than to the emotional changes arising from the doctor–patient relationship. This is an interesting book written by an individualist who is really an eclectic in matters of psychodynamics, but has evolved a language and a technique of his own.

E. STENGEL

The Chemistry of Thinking. By G. HUMPHREY and R. V. COXON. Springfield, Ill.: Thomas. 1963. Pp. xii+148. $6.78.

This book is made up of two sections. The first is written by the psychologist of the team and is a brief survey of some aspects of thinking. The expert will find it inadequate, while the beginner will be puzzled by it. The physiological section is a brief well-written account of the more important aspects of biochemistry relevant to the functioning of the nervous system, but it has a very tenuous connexion with thinking. However, there is one chapter devoted to the neurophysiological problems involved in learning, and this ends with a short account of DNA and a brief note on Hydén's RNA theory of learning. There is a final chapter on psychopharmacology in which it soon becomes apparent that the author is dealing with a subject which is outside his specialist field.

To call this book *The Chemistry of Thinking* is a misuse of words. Anyone interested in the neurophysiological processes which may be related to mental activity will get no help from this book. The second edition of H. W. Magoun's *The Waking Brain*, which has recently been issued by the publishers of this book, is easily the best up-to-date review of this problem.

FRANK FISH

Probation and Mental Treatment. By MAX GRÜNHUT. London: Stevens. 1963. Pp. vii + 56. 15*s*.

Crime and the Social Structures. By JOHN B. MAYS. London: Faber and Faber. 1963. Pp. 256. 30*s*.

Crime and the Criminal Law. By BARBARA WOOTTON. London: Stevens. 1963. Pp. viii + 118. 21*s*.

During 1963, an inconspicuous, though rich harvest of criminological books was reaped, among which the above three have their distinct place.

Dr Grünhut analyses, straightforwardly, a sample of 636 of the 882 offenders who in 1953 were dealt with by the courts under section 4 of the Criminal Justice Act, 1958, requiring the offender to undergo medical treatment in conjunction with probation up to 12 months. He found that 275 offences were committed against property, 216 were sexual, 65 suicides, 32 violence, 48 miscellaneous. Classification of the medical reports of probationers revealed 195 psychopaths, 28 of low intelligence, 58 schizophrenics, 52 depressives, 38 anxiety neurotics, 36 physically ill. The majority received psychotherapy, to which exhibitionists responded well and psychopaths badly. Encouragingly, out of every 100 treated offenders, 70 were discharged with a favourable prognosis, 56 of whom were not reconvicted within 12 months after termination of probation and 13 out of 30 with an unfavourable medical prognosis proved to be non-recidivists. Although neither 'depth' nor a factor analysis is attempted the trends are clear and interesting.

Dr May's very readable book is based on the hypothesis that society itself creates crime as a necessary corollary to progress. He offers many useful suggestions as to how the inevitable crime rate might be kept down within reason: by recommending that society should face up to its responsibility to adolescents through an extended and integrated use of youth clubs; Police Juvenile Liaison officers; the recognition by teachers that their function as educators extends into the field of social science; and that the Ministry of Education should be partly concerned through education with crime prevention.

Lady Wootton pulls no punches. Essentially she pleads that judges and magistrates should follow a scientific sentencing policy based eventually on statistical prediction studies in order to prevent recidivism rather than relying on the old penal concept of assessing emotionally the offender's measure of guilt and moral responsibility. In general, she favours indeterminate sentences and the European inquisitorial (rather than the British accusatorial) system with certain safeguards. She recommends that courts of the future should, after establishing guilt and deciding on institutionalization, allow non-judiciary Boards to handle the details of sentences, and that judges receive instructions in scientific penology. She also reviews elusive crime-causation theories and calls for a solution to the question of why the crime rate is appreciably lower for women than men. Her provocative and scholarly analysis of the inconsistencies of the law in relation to mentally abnormal offenders is, together with the rest of her book, most refreshing.

All three authors are courageously outspoken in favour of scientifically based research and action in criminology.

R. G. ANDRY

Vol. 55. Part 3 August 1964

THE
BRITISH JOURNAL
OF
PSYCHOLOGY

EDITED BY
BORIS SEMEONOFF

AIDED IN THE SELECTION OF PAPERS BY
C. BURT
AND

F. C. BARTLETT
E. A. BOTT
D. E. BROADBENT
R. S. CREED
J. DREVER
G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
M. D. VERNON

ASSISTANT EDITOR
R. M. McKENZIE

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W. 1
AMERICAN BRANCH, 32 EAST 57TH STREET, NEW YORK 22, N.Y.

Price Twenty-five Shillings net (*U.S.A.* $4.25)

Preparation of Papers for the British Journal of Psychology

The *British Journal of Psychology* subscribes to the *Fair Copying Declaration* of the Royal Society under the terms of which reproduction of material published in this Journal is permitted.

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. All papers, including reviews, must be typed (double spacing) on one side of the paper only and the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of most articles published from Vol. **47** (1956) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

2. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs etc. is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered caption of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following example:

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.
RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* **50**, 83–96.

Note that World List abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

Brit. J. Psychol. (1964), 55, 3, pp. 249–251
Printed in Great Britain

OBITUARY NOTICE

ERNST KRETSCHMER

I first met Dr Ernst Kretschmer, Professor of Psychiatry and Neurology in Marburg, at the Maudsley Hospital, shortly after the Second World War. He was a short, dapper man, pyknic in body-build, and cyclothymic in character; well dressed, prosperous and looking more like a businessman than a professor. What struck me most, however, was that he was middle-aged; in my mind I had vaguely grouped him with Wundt and Kraepelin as among the early pioneers of psychology and psychiatry and had assumed that he too had handed over to the next generation. The error was perhaps permissible. *Körperbau und Charakter*, the book which brought Kretschmer world-wide fame, was published in 1921 when the author was just 33 years old and was Privatdocent and assistant to Gaupp in Tübingen. This, however, was not his first book; Kretschmer had been known in professional circles since 1918 when he brought out his book about *Sensitiven Beziehungswahn*. This early book contained a protest and a reaction against Kraepelin's typology and a plea for multi-dimensional types of diagnosis; it is ironic that his life's work thereafter was destined to derive from, and in turn buttress, a strict Kraepelinian typology.

Other books by Kretschmer in his 30's were his book on *Medizinische Psychologie* which appeared in 1922, and his *Hysterie* which appeared the year after. At the end of the 1920's he published his great work *Geniale Menschen* (1929) in which he applied his general typological theories to the problem of genius, grouping his chosen sample into cyclothymes and schizothymes, relating their performance to their body-build and generally attempting to demonstrate the detailed application of his principles to these highly gifted people.

As a psychiatrist and writer, Kretschmer had certain gifts which may account for the fact that for many years he was regarded as the foremost German medical psychologist. In the first place he was a great writer, clear, evocative and able to pin down the description of a character in a few supremely well-chosen words. This gift is of particular importance in classical German psychiatry, which lays great stress on case histories and description generally. The importance of this gift is even more apparent for a writer who tries to appeal not only to fellow psychiatrists and psychologists but who wants to interest the intelligent lay public in his work, as Kretschmer tried to do in his book on genius. You might disagree with what Kretschmer had to say, but you were never in doubt as to what it was that he was saying, and you would always enjoy his graceful way of saying it.

Kretschmer's second great quality was tolerance. When it was fashionable to decry Freud for reasons which had nothing to do with the truth or falsity of his notions, Kretschmer refused to take part in this denigration of new theories and hypotheses, instead incorporating what he considered true and valuable in his teaching. When clinical psychology began to appear on the scene, Kretschmer for a time was the only psychiatrist who allowed psychologists to work in his clinic and take part in his teaching rounds. These are only a few examples of what was an outstanding feature of the man, and his great influence on German psychiatry was in large effect due to this ability to see good in many systems, to select contributions of substance from all sides and to refuse resolutely to wear blinkers and condemn unheard any school.

Psychologists of course are concerned in the main not with his psychiatric contributions but with his work as a typologist. The notion of 'type' was from the beginning a central one in Kretschmer's thinking. He considered it

> the most important fundamental concept of all psychology. Nature...does not work with sharp contrasts and precise definitions, which derive from our own thought and our own need for comprehension. In nature, fluid transitions are the rule, but it would not be true to say that, in this infinite sea of fluid empirical forms, nothing clear and objective could be seen; quite on the contrary. In certain fields, groupings arise which we encounter again and again; when we study them objectively, we realize that we are dealing here with focal points of frequently occurring groups of characteristics, concentrations of correlated traits....What is essential in biology, as in clinical medicine, is not a single correlation but groups of correlations; only those lead to the innermost connexions....What we call, mathematically, focal points of statistical correlations, we call, in more descriptive prose, constitutional types....A true type can be recognized by the fact that it leads to ever more connexions of biological importance....

Thus Kretschmer, and it will be clear that his point of view has nothing to do with the typical American textbook simplification which represents types as being 'either/or' principles of classification. Normal distributions are implicit in Kretschmer's thinking and when I asked him about his opinion on this point he emphatically agreed that there was nothing in his writing or thinking which could encourage any other belief. He was rather contemptuous, perhaps rightly so, of much American work that had been done to investigate his hypotheses; he said that the work might have been better done if the experimenters had first read his book!

In part, there is here perhaps a genuine misunderstanding. Kretschmer kept working at his theory throughout his life, and, as is the custom in Germany, incorporated more recent studies in the later editions of his book, which ran into over twenty editions by the time of his death. English-speaking writers and experimentalists usually relied on the English translation of a quite early edition which was by no means up-to-date and had not reflected a great deal of the work that was done by Kretschmer and his school. Nevertheless, this is little excuse for the manifest ignorance of many authors who wrote about Kretschmer without realizing the ways in which his teaching had grown and changed. Kretschmer became even more disenchanted with American psychology when Sheldon essentially took over his teaching, furbished it up with a few esoteric terms and improbable hypotheses and offered it as new and original to a profession which did not have the knowledge or the historical sense to see this substitution for what it was.

With all his superb gifts there was one that was lacking, and this lack probably proved fatal for Kretschmer as a genuine innovator in human typology. Kretschmer was 'literate' in C. P. Snow's phrase but he was not 'numerate'. In the quotation given above Kretschmer keeps talking about correlations but he never in fact calculated any correlations. He relied essentially on insight rather than on calculation and thus fell short of the essentials of scientific proof. Even in his studies of body-build he was suspicious of any index calculated in terms of body measurements, although sometimes contemptuously permitting lesser mortals to use such crutches.

When in later years critical experimental studies, such as those by Payne and Brengelmann, showed an absence of correlation where he had predicted substantial relationships, he was surprised and hurt. Yet there had been other clouds on the horizon which should have given him pause. There are obvious dangers in choosing schizophrenia and manic-depressive insanity as the extreme ends of a personality dimension supposed to encompass normal people as well, when genetic studies leave little doubt that these disorders are strongly determined by the action of probably single genes. Again the relationship between body-build and insanity predicted by Kretschmer does appear to be a reality but it is so tenuous as to be practically useless and incapable of supporting the type of experimental research encouraged by Kretschmer, which uses body-build as a kind of scaffolding or link.

This lack of ability to see things in a quantitative fashion is not perhaps surprising in one who was brought up in a climate of phenomenological psychology; it is all the more to be regretted, however, when it is realized that Kretschmer was not only a brilliant and intuitive psychiatrist but also a highly gifted experimentalist. He had the very rare ability to postulate psychological traits and then construct experimental tests embodying these traits; I know of no one in the personality field whose ideas in this connexion were superior to Kretschmer's. The later editions of *Körperbau und Charakter* are gold mines for Ph.D. students in search of a thesis; if only psychologists could be called away from their worship of the Rorschach to an investigation of the brilliant ideas pioneered by Kretschmer, how much more quickly would psychology advance!

H. J. EYSENCK

Brit. J. Psychol. (1964), 55, 3, *pp.* 253–294
Printed in Great Britain

PERSONALITY VARIABLES IN PSYCHIATRIC CLASSIFICATION: A SYMPOSIUM

This symposium was held during the Annual Conference of the British Psychological Society at Reading in March 1963. It was inevitable that much of it would be concerned with 'Eysenckian' ideas. We were convinced that extreme positions were not appropriate in the present stage of development of the subject. The aim, therefore, was to invite people who were doing independent work on different aspects of personality and psychiatric diagnosis, and who would present their views with due concern for scientific standards. Their papers are reproduced here, substantially as they were delivered, with only minor alterations. References are consolidated at the end of the symposium.

Four of the papers put forward points of view which differ in one or more respects from Eysenck's. Patrick Slater's contribution is a reassessment of the basic methodology of factor analysis and other forms of multivariate analysis which have played an important role in this field. Eliot Slater considers Eysenck's bi-dimensional description of neurosis in the light of evidence from psychiatric genetics. He discusses whether such evidence demands a pluridimensional concept. Foulds draws attention to a distinction between relatively enduring personality variables on the one hand and symptoms and signs of illness which are relatively transient on the other. He also argues that it is reasonable to define a 'personal illness' continuum, along which both neurotic and psychotic groups may be discriminated from normal people. The psychotics occupy a more extreme position on the continuum than neurotics. Ingham and Robinson discuss results obtained with the Maudsley Personality Inventory and, in particular, the mean scores of various groups of dysthymic and hysteric patients. They question the practice of regarding hysterics as a homogeneous group and argue that the characteristics of hysterics as observed by questionnaire responses depend upon the clinical features accorded greatest weight in diagnosis. Eysenck states his present position and summarizes some of the experimental work that his ideas have helped to engender. He argues that we are now in a much better position for the development of a scientifically based system of classification than we were a few years ago.

J. G. Ingham
Convener

SPECIFIC FACTORS AND MINOR COMPONENTS

By PATRICK SLATER
Institute of Psychiatry, the Maudsley Hospital

INTRODUCTION

This paper is in two parts. The first submits an argument that correlations between tests provide no evidence of the existence of specific factors, but only of components which are non-specific, even when they are minor ones. The second provides evidence that minor components can often be of psychological interest.

It was through considering evidence of the kind given in the second part that I was led into formulating the argument presented in the first. But when I try to persuade others to take an interest in minor components I find them often impervious to the idea because they are accustomed to assuming that beyond a few common factors with large variances nothing is to be found by factor analysis but negligible specifics. So in this paper the argument precedes the evidence. I hope that any reader who finds the argument difficult to follow or misleadingly expressed or suspects there is a fallacy in it will proceed to the second part; and then, if he finds the evidence there adequate and convincing, he will be persuaded that even if the argument as stated is not correct there must be some form in which it can be stated correctly.

It is a very general one, not restricted to a particular method of factor analysis nor to a particular type of test material, and it continues to apply when the tests under consideration are tests for assessing personality variables to be used in psychiatric classification, as the examples from Eysenck (1960*d*), Franks *et al.* (1960), Hope (1963) and Slater (1944) illustrate.

I. THEORETICAL CONSIDERATIONS

Most methods of factor analysis (applied to correlations between psychological tests) assume the existence of as many specific factors as there are tests. Spearman's two-factor theory (Spearman, 1923) was the first to do so, and incorporates the assumption in its simplest form, so it is the most convenient one to examine in detail.

The theory asserts that all cognitive operations are served alike by a general ability, and each particular operation is also served by a specific ability peculiar to it. Though it is not a theory restricted to properties of psychological tests it applies to test scores and leads to definite expectations about their variances and correlations. It may be stated with reference to test scores as follows:

> *Hypothesis*. Scores on a given cognitive test, t depend on two factors only: g, a general factor which serves performance on all cognitive tests, and s, a specific factor which serves performance on t only. The fundamental equation
>
> $$t_i = g_i + s_i$$
>
> relates t_i, an individual's score on the test, to g_i and s_i the measures of his general ability and his special ability on standard scales.
>
> *Two derived equations* define the factor loadings r_{tg} and r_{ts}. The first concerns the self-correlation of test t, that is to say its normalized variance, which must equal unity
>
> $$r_{tt} = 1 = r_{tg}^2 + r_{ts}^2. \quad (1)$$

This asserts that it can be resolved into the sum of the squares of two quantities r_{tg}, the g-loading of the test, and r_{ts}, its specific factor loading. The second concerns the intercorrelations of a pair of tests, 1, 2:

$$r_{12} = r_{1g} . r_{2g}. \tag{2}$$

It asserts that their correlation depends exclusively on their g-loadings as they have no factor but g in common.

The logical status of the expressions is as follows: g and s are hypothetical variables which cannot be isolated *experimentally* or observed or measured *directly*. The equations are deduced from the theory. They must hold good within the limits of experimental error if the theory is correct; but they are not axiomatic and cannot be considered proved unless some independent proof can be given for the theory from which they are deduced. If some proposition deducible from them is found to be false they become suspect and so does the theory. A factor loading is the inferred estimate of the correlation between a test and one of the hypothetical variables.

When factor analyses have been made by this method interest has focused mainly on the g-loadings of the tests. Typical conclusions are that one test measures g better than another because it has a higher g-loading, or that g must have certain psychological properties because the tests with similar loadings have certain psychological resemblances.

The results have seldom been used to estimate the general intelligence of individual subjects, and never, I believe, to estimate their special abilities. There are no theoretical attractions to the latter procedure. Of what use are estimates of specific abilities when according to the theory they do not predict anything beyond the scores from which they were derived?

Still, we may consider how these estimates could be derived and note what their properties should be according to the theory. There would be a certain number of tests in the experiment, say n. Their g-loadings could be found and thence a set of weights. The g-score of an individual, that is the estimate of his general intelligence, would be obtained as the weighted sum of his scores on the n tests. Then his s-score for each test in turn could be obtained by substitution in the fundamental equation. There would be n of these scores and the theory asserts that they should be independent of one another.

Thurstone's restatement (1947) of equation (1) retains its form in outline. A collective term, the communality of the test, replaces r^2_{tg}; and the final term r^2_{ts}, renamed the test's specificity or uniqueness, is sometimes interpreted if not treated statistically as a composite also. Since then there has been controversy; doubts about the legitimacy of supposing *that every single test has one or more specifics which affect its variance while leaving its covariance with every other test unaffected* are probably among the main reasons for the 'coolly negative attitude' most statisticians outside psychology have towards factor analysis (Kendall, 1950; Ehrenberg, 1962). But most psychologists follow procedures incorporating this assumption uncritically. Controversy has produced no more convincing justification than Spearman's for reducing the variances of the variables to an extent at choice when factorizing, without altering their covariances. Cattell (1952) accepts the postulate. Lawley & Maxwell (1962) describe it as the basic assumption of factor analysis. Accordingly, they exclude principal component analysis, which does not depend upon it, from their survey. I shall maintain their distinction and contrast factors and factor analysis on the one hand with components and component analysis on the other.

Factor analysis, so defined, usually terminates after a few common factors have been extracted and the factor loadings have been obtained. Then the interpretation begins. But again it is possible to calculate a subject's factor score for each of the common factors, and finally for each of the specific factors. The n tests will yield n specific factor scores as before and the theory maintains that they should be uncorrelated.

But they will turn out to be correlated. In fact, no variable with the properties ascribed to specific factors can be derived from test scores. They are not there. To take the simplest conceivable case, let us suppose two tests known to have the same g-loadings are given to an individual, and his scores expressed in standard units are found to be a on test 1 and b on test 2. Then his g-score is simply the average of the standard scores, that is $(a+b)/2$.

The estimate of his special ability on each test is the difference between his score on that test and his g-score. So his 'specific' score on test 1 is $(a-b)/2$, and on test 2 it is $(b-a)/2$. The two 'specific' scores are equal to one another but opposite in sign. And if we calculate specific scores for a number of individuals we shall find that the supposed special abilities on the two tests have a correlation of -1, not a zero correlation as the theories require. The same factor has been estimated twice over, once in reverse. And the estimate is a weighted sum of the standard scores on both tests, so the factor is not specific to either test but common to both.

The same result, i.e. a correlation of -1 between the two so-called specifics, is obtained when the g-loadings of the two tests are not equal. And it can be generalized for all larger numbers of tests and factors (Ehrenberg, 1962). If a test has a significant multiple correlation with the other tests included in the experiment, then the whole of its variance can be referred to common factors. So it is safe to assume that not even one specific factor will be found in any experiment where the tests all have some significant correlations. All the factors will be common factors. This must be true for any method that depends on summation.

The following argument is quite general: Suppose that N persons are given n tests and the results are set out in a table X, the transpose of the usual form (following Kestelman's 1952 notation), with a row for each test and a column for each person. The *test score* on test t is a variable x_t, of which N values are observed. The *factor score* for factor k (which may be common or specific) is defined by a set of weights $b_{k1}, b_{k2}, \ldots, b_{kn} = \mathbf{b}_k$, to be applied to the test scores of any individual, so the formula

$$f_k = b_{k1}x_1 + b_{k2}x_2 + \ldots + b_{kn}x_n \quad \text{or} \quad \mathbf{b}_k\mathbf{X} \tag{3}$$

will give the factor scores for the N individuals in factor k.

We may begin by choosing a set of weights $\mathbf{b}_1$ for factor 1, then another $\mathbf{b}_2$ for factor 2, and so on, in such a way that f_k is not a linear dependent function of $f_1, f_2, \ldots, f_{k-1}$, up to a maximum possible number n. But we cannot proceed beyond this point because the n equations (3) can then be solved for the x's, giving n equations of the form

$$x_t = c_{t1}f_1 + c_{t2}f_2 + \ldots + c_{tn}f_n$$

(when the c's are computed constants); and these can be substituted into equation (3) for f_{n+1}. Then f_{n+1} is expressed as a dependent function of $f_1, f_2, \dots, f_n$.

This argument does not require the factors to be set down in any particular order. So by rearrangements any one of them may be taken as f_{n+1}, and shown to be dependent on not more than n of the others. It only requires that the number of factors extracted, including the so-called specifics, exceeds the number of tests. It is true exactly, not only approximately; true whether the weights are normalized or not, whether the factors are orthogonal or oblique, whether the test scores are standardized, normally distributed, etc., or not; true in the sample and in the population. Moreover, the argument is one that can be extended to non-linear dependent functions.

For example, Thomson (1951), who discusses the estimation of a man's factors with great care and detail, obtains equations for estimating g and four specifics from the intercorrelations of four tests z_1, z_2, z_3, z_4, viz:

$$\hat{g} = 0{\cdot}553z_1 + 0{\cdot}259z_2 + 0{\cdot}160z_3 + 0{\cdot}109z_4,$$

$$\hat{s}_1 = 0{\cdot}737z_1 + 1{\cdot}313z_2 - 0{\cdot}215z_3 - 0{\cdot}145z_4,$$

etc. Evidently the four equations for the specifics can be used to compute values of c and obtain expressions of the form

$$z = c_1\hat{s}_1 + c_2\hat{s}_2 + c_3\hat{s}_3 + c_4\hat{s}_4$$

for the z's. And when these are substituted in the equation for $\hat{g}$, $\hat{g}$ will have been expressed as an exactly dependent function of the four specifics. Similarly, $\hat{s}_1$ can be expressed as a dependent function of $\hat{g}$, $\hat{s}_2$, $\hat{s}_3$ and $\hat{s}_4$, etc. We have no need to inquire what persons were tested, what tests were used or what method of analysis was applied in order to satisfy ourselves of this.

When a factor analysis alleges that certain factors such as g, s_1, etc., have been found we may inquire what the subjects' measurements are in all these variables. The answer to this question must either be an inconclusive one admitting an indefinitely large number of different possibilities (Kestelman, 1952) or a precise one giving formulae for obtaining estimates, such as Thomson's for $\hat{g}, \hat{s}_1, \dots, \hat{s}_4$. Now, if it is reasonable to hold that measuring n variables in N cases ($n < N$) does not show that any variable has occurred in the experiment unless measurements of it have been obtained, we may say that the occurrence of *all* the factors has not been proved, and that the only variables found are the tests themselves and their derived functions, such as $\hat{g}$, $\hat{s}_1$, etc., of which not more than n can be independent of one another. Thus the factorist can be required to sacrifice his specifics to retain his common factors.

A diagram may help to explain the argument. The ellipse in Fig. 1 illustrates a scatter of standardized scores on a pair of tests, 1 and 2, which are correlated. It is because of the correlation that the scatter is elliptical in form and lies diagonally across the reference axes, 1 and 2.

The sum of a pair of scores weighted equally defines a point on the P-axis; so that is where interest focuses when equal weights are used. It is the axis of the major component, the dimension where variation is widest. After the variation there is partialled out the variance left is confined to Q, the axis at right-angles to P. It is the axis of the minor component, where variation is most restricted. And it must intersect the reference axis, just as the P-axis does, so measures of Q are also weighted sums of the test scores.*

Suppose a common factor is extracted by some method of factor analysis that does not coincide with component analysis. Then interest becomes focused on some other axis, like the F-axis, which lies obliquely across the reference axes but diverges from the P-axis to some extent. Still the residual variance must be confined to one other axis, G, because the dispersion is only two-dimensional. And it is an oblique axis because it is at right-angles to the F-axis and the F-axis is oblique. Thus if a common factor is selected for consideration what is left is another common factor. No specific factors will have been found. If there are more than two tests, consider them two at a time: the argument must hold.

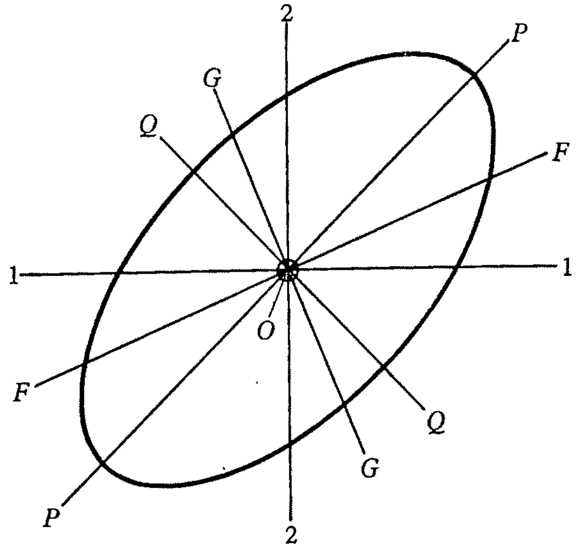


Fig. 1. Alternative co-ordinate systems for a bivariate frequency distribution. $O1$, $O2$, the reference axes; OP, OQ, the major and minor components; OF, OG, another pair of axes (OF oblique to $O1$, $O2$; OG orthogonal to OF).

In general, suppose any number of independent tests, say n, have been used, and a certain number of common factors, say c, have been extracted, then the analysis separates the n space of the reference axes into two subspaces, one of c dimensions where the variance is relatively large and one of $n-c$ dimensions where it is relatively small, both oblique to the reference axes. The former is commonly treated as the region of psychological interest and the latter is disregarded.

If the assumption that there is a specific factor for each test is dropped, the statistical theory may be modified by introducing the assumption that the minor sources of variation occurring in all experiments are liable to affect all dimensions indiscriminately. Then if the design of a particular experiment extends the variation effectively in a limited number of dimensions they can be isolated by a principal component analysis, and the residual variance should be homogeneous and measure the extent of experimental error. M.S. Bartlett and his colleagues (Bartlett, 1948, 1950; Kshirsagar, 1961) have worked out many of the mathematical implications of this modified theory.

---

* Components in general are defined by their latent vectors, which specify their axes, and their latent roots, which measure the extent of variation along their axes.

It can also be tested experimentally by locating the minor dimensions where the variance is supposedly at random, and examining their contents. My investigations of these dimensions have led me to the conclusion that this modified theory is an oversimplification. It is not unusual to find that minor dimensions are of psychological importance.

Fig. 1 may help to show how essential it is to use a principal component analysis to locate them. A slight displacement of the major axis P will also displace the next axis Q and enlarge the residual variance. A succession of such displacements in the course of an analysis will have cumulative effects. When the minor components are reached their magnitude will be greatly exaggerated and the estimates of their direction will be far out. The results will be disastrous when the components requiring intensive investigation are the last to be found.

II. EXPERIMENTAL EVIDENCE

There are no general grounds in science for assuming that dimensions where variation is restricted are uninteresting. On the contrary, science is mainly a search for uniformities and invariances in nature. Consider body temperature, for instance, which is almost completely invariant among normal people. Yet it is so reliable a measure of what might be called a general factor of feverishness that housewives can be advised to keep a thermometer in the house and send for the doctor when a child has a temperature. On the other hand, pulse rate, blood pressure, respiration rate, etc., are far less reliable measures for elementary diagnostic purposes because they vary *too* widely among normal people.

For a simple example from psychological testing take the results shown in Fig. 2. They come from giving two vocabulary tests, M and F, to a group of boys and a group of girls of the same age in a mixed secondary school (Slater, 1944). In each group the children tend either to do well on both tests or badly on both. Scores are correlated positively to very nearly the same extent and the standard deviations are approximately equal. Factor-analytic theory would suggest that the scores depend on verbal ability, best measured by adding them together with equal weights. The residual variance would not be considered worth investigating because the theory attributes it to specifics, errors, etc. But in point of fact the minor dimension of variance within the groups, measured by the differences between the M and F scores, coincides approximately with the dimension which discriminates most effectively between the groups. The data show only two components, and both are of great psychological interest, though naturally from different points of view.

In a more elaborate experiment (Eysenck, 1960*d*, vol. 2, chap. 8) ten tests were given to neurotic patients of six different classes and a sample of normals, to find how effectively they discriminated between the groups. Table 1 lists the components according to the size of their latent roots and shows which were found to contribute effectively. The two largest components are included, but so are the three smallest and two of the five intermediate ones. Thus there is no consistent relationship between the amount of individual variation in a particular dimension and its effectiveness for discrimination. No matter how small the variation may be, the dimension need not be devoid of psychological interest. And conversely evidence that the variation is relatively large may not imply that the dimension is important.

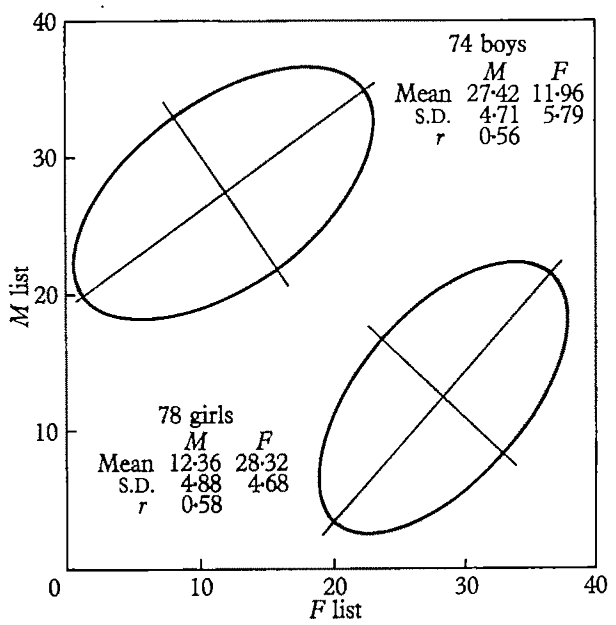


Fig. 2

Table 1. *The principal components of ten tests in order of magnitude, showing which discriminated effectively between seven diagnostic classes*

Component	Proportional magnitude (%)	Significance of discrimination
First	29·9	**
Second	16·3	**
Third	11·6	—
Fourth	9·5	**
Fifth	8·6	—
Sixth	6·5	**
Seventh	6·1	—
Eighth	5·2	**
Ninth	3·8	**
Tenth	2·5	*

* Significant at the 5 % level. ** Significant at the 1 % level.

Rather similar results have been published recently by Jeffers (1962). He obtained measurements of 10 variables for comparing trees of the same species grown from seed from 13 different sources; and extracted 5 components with proportional magnitudes ranging from 50·5 % for the first to 4·6 % for the fifth. These two components were the ones that revealed significant differences between the sources; the intermediate ones did not.

Another way of comparing major and minor components is in terms of their test–retest reliability. Some relevant results have been obtained from the preferences expressed for 12 objects by 400 people on two occasions (Slater, 1960*a*). Two dispersions are formed in the reference space, giving two vectors for each component. A component may be considered reliable or stable if its vectors tend to concide. A suitable statistic for measuring the tendency is the cosine of the angle between them, which varies from 1 when the vectors coincide exactly to zero when they are at an angle of 90 degrees, and is equivalent in many respects to a correlation.

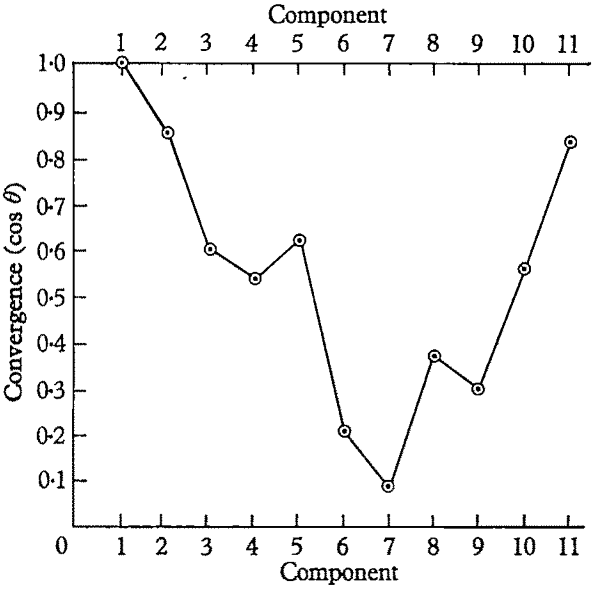


Fig. 3. The test-retest convergence of the 11 non-zero components found in an analysis of personal preferences.

Fig. 3 shows the comparison. If the theory is correct, and the latent roots are more saturated with error the smaller they are, we should expect to find the axes diverging farther and the cosines falling steadily as we pass from the first component to the last. As far as the seventh the cosines fall in accordance with this expectation. But instead of levelling out there they rise again with a fine disregard for theory. The smallest non-zero component is just about as stable as the second. The cosines are: second, 0·864; eleventh, 0·838.

The stability of major and minor components under repeated sampling can also be compared. Since an earlier draft of this paper was circulated results of this kind have been referred to me by Hope (1963). He used questionnaire tests of five types of hostility, such as projective hostility, guilt and criticism, and obtained results from a sample of 47 normals and 169 neurotics, cf. Table 2. Table 2 (*b*) shows that the largest and the smallest components have the greatest stability. The third and fourth components, which have approximately the same size in the Neurotic sample, appear in reverse among the Normals; but there is distinct evidence of stability throughout the series. In Hope's opinion there is a reasonable interpretation for the fifth component.

Table 2. *Principal components of five assessments of hostility*

(*a*) The proportionate magnitude of their latent roots.

	Per cent among	
Component	Neurotics	Normals
First	52·3	51·9
Second	21·2	16·1
Third	10·4	13·2
Fourth	9·8	10·4
Fifth	6·3	8·4

(*b*) The cosines of their vectors.

		Components from the normal sample				
		1	2	3	4	5
Components from the Neurotic sample	1	**0·99**	0·06	−0·06	−0·01	0·09
	2	0·08	**−0·79**	0·54	−0·27	−0·13
	3	−0·02	0·34	0·02	**−0·93**	−0·12
	4	0·01	0·50	**0·84**	0·19	0·10
	5	−0·08	−0·12	−0·00	−0·16	**0·98**

The largest cosine in each row and column in bold.

Other evidence of the stability of a minor component is shown when the direction of its vector is unaffected by the introduction of additional tests. The data of Franks *et al.* (1960), discussed below, are worth examining from this point of view.

If the variance of a minor component is not due to error it should be controllable to some extent. And indeed it is possible to exclude variance from a dimension entirely or restrict it to any desired extent. These are operations that can be performed either deliberately or inadvertently, and either mathematically or experimentally.

The basic rule for complete exclusion is a formal mathematical one: a dispersion is restricted to $n-1$ dimensions in an n-space if the n variables include one which is a linear dependent function of any of the others. A factor score, for instance, is such a variable. A set of variables consisting of $n-1$ tests and a factor score derived from them constitutes an $n-1$ dimensional dispersion in an n-space with one empty dimension. This is a lemma to the proposition that a factor score does not measure any variable beyond those from which it is calculated nor prove that any further variable exists.

The formal rule for complete exclusion is employed in the computing procedures for partialling out a variable or a component. Thus an $n \times n$ table of residuals after one component has been partialled out provides another example of a dispersion in an n-space with one empty dimension.

To achieve the same result experimentally the design of the experiment must ensure that the variables measured are connected by some linear relationship. One such relationship, often incorporated in psychological experiments, is that the sum of the measurements is the same in every case. Suppose, for instance, that subjects are put for a fixed length of time in an environment which allows them certain mutually exclusive kinds of behaviour; and the records kept are of the amount of time a subject actually spends doing A, B, C, etc. Then there will be a linear relationship between the times because their total is the same for every subject. Bindra and others (e.g. Bindra & Spinner, 1958) have published much material of this kind but have not applied multivariate analysis to it.

Or suppose A, B, C, etc. are a set of objects a subject has to choose between. If he chooses A in preference to B the entry in his record is 1 for A and 0 for B, and so on for each choice he makes. As every subject has the same number of choices the total of his entries must be the same, and again the linear relationship occurs.

Table 3 shows a set of correlations derived from data of this kind. It provides an interesting contrast to tables of correlations between cognitive tests, where the entries are predominantly positive. There is not one positive correlation here.

Table 3. *Correlations between preferences for the perfumes of five soaps (25 informants)*

	Perfume			
	B	*C*	*D*	*E*
A. Sandalwood	−0·46	−0·35	−0·23	−0·09
B. Blue carnation	—	−0·29	−0·18	−0·16
C. Lily of the valley	—	—	−0·01	−0·23
D. Flowers of lavender	—	—	—	−0·47
E. Rose geranium	—	—	—	—

A general tendency to negative correlation arises in all tables where the rule applies that the measurements correlated have the same total in every case. Such tables present awkward problems for factor analysis, but component analysis handles them without difficulty (Slater, 1960*a*) and concludes correctly with a last latent root of zero.* Experiments which incorporate this rule have a common component. Moreover, they are generally similar in their psychological content, being experiments in which a number of objects or courses of action, etc., are offered and the subject is required to make a choice. Thus there might be a temptation to associate the component with the content and give it a name, say a factor of 'choice' or 'preference' or 'disjunctive judgement'. But the fallacy of such reasoning is evident in this context because there can be no doubt about how the component is controlled here.

When the rule is enforced rigorously a component with zero variance is obtained; by relaxing it, but only slightly, a component with exceptionally small variance can be obtained. An example of such a component was published in 1953 (Slater, 1953), and I shall not reproduce it here.

Just as there are rules for combining tests in an experiment to reduce or exclude the variance in a particular dimension there are also rules for producing or increasing it. There is no need to illustrate rules of the latter kind: the assertion that a factor exists implies that a rule can be formulated for producing it, even when no rule is actually stated. Thus we may find rules for producing particular factors, rules for producing any number of factors of a particular kind, and so on.

There may also be rules for controlling variation in several dimensions simultaneously. This is a possibility that may be worth more systematic investigation. For instance there may be ways of producing variation in one dimension which imply excluding it from another. My final example seems to be of this kind: in an experiment designed to produce variation in two dimensions variation in a third has been restricted.

---

* The characteristic vector of this component is $\mathbf{1} = (1, 1, \ldots, 1)$, with the property $\mathbf{V}\,\mathbf{1}' = 0\,\mathbf{1}'$ where $\mathbf{V}$ is the variance–covariance matrix of the variables.

The data come from a paper published by Franks *et al.* (1960). The D, C and R scales of Guilford's S.T.D.C.R. inventory were used to provide indices of neuroticism and introversion: extraversion supposed to explain the factorial composition of four scales from the Minnesota Multiphasic Personality Inventory. There is no need to quote the whole 7 × 7 table of correlations. The correlations between the three Guilford scales are shown in Table 4(*a*). The high correlation between the first two variables, and the difference in sign between the remaining correlations, make this little 3 × 3 table another that is rather awkward for factor analysis, but again component analysis can be applied without difficulty.

Table 4

(*a*) The intercorrelations of the D, C and R scales of Guilford's S.T.D.C.R. inventory found in a sample of 200 normal adults.

Scale	C	R
D	0·9200	−0·1924
C	—	0·0860

(*b*) Their principal components.

		First	Second	Third
Latent root		1·9285	1·0271	0·0444
Normalized	D	0·7106	−0·0663	0·7007
Vector	C	0·6970	0·2024	−0·6877
	R	−0·0962	0·9771	0·1898

The relative size of the latent roots in (*b*) indicates that although variation is recorded in three dimensions it is practically confined to two. Over 98 % is taken up by the roots of the first two components. The authors' interest is confined to these two: they analysed the complete 7 × 7 table including the four MMPI scales by Lawley's method; made a rotation; and finally selected an axis for neuroticism which coincides almost exactly with the first component of the 3 × 3 table, and an axis for introversion: extraversion that coincides almost exactly with the second. The cosines are 0·9994 and 0·9962.

When the components of the 7 × 7 table are extracted the last is found to have an exceptionally small latent root, less than a quarter of the preceding one. Its vector coincides almost exactly with the smallest component of the 3 × 3 table. The cosine is 0·9994. Thus the tests which locate the major dimensions of interest to the authors also define another unnoticed dimension where variance is restricted. This may explain why their tests of significance break down. They found significant residuals after the extraction of three factors although they were only able to interpret two. Bartlett's test would have shown all six preceding components significantly greater than the last.

This example may be typical of many. It is quite usual to include reference tests for the factors of greatest interest in a factor-analytic study, that is to say, sets of tests which correlate highly with one another but not with the tests in other sets. Such a procedure seems liable to restrict variance in some dimensions while it extends variance in others. The question may then arise whether any lawful relationships between the variables may not be defined as well by the dimensions of maximum invariances as by those of maximum variance. A complete analysis will provide expressions of both kinds. Indeed there are many conditions where a number of different expressions are found to be equivalent, and the presence of a specially small latent root in a matrix has many curious implications. A relevant example is given by Kendall (1957, Ex. 5, 1) in a discussion of the interpretation of a multiple regression equation.

Another question is whether tests available at present for the significance of factors and components rest on trustworthy assumptions. It seems difficult to reconcile any current notions of the significance of factors with the evidence I have given to show that the variances of minor components are not properly attributable to specific abilities or experimental errors but may be stable and reliable and have predictive value, and can be brought under experimental control.

If systematic studies are made of the minor components appearing in experiments where their appearance is not methodically controlled, I would expect the results to show that there are some, perhaps of a definable kind or kinds, where no specially small components are to be found; in some they are introduced by trivial or arbitrary restrictions, such as scoring procedures or duplications of test material; and finally that in some they have features of real psychological interest. I would expect studies of methods for controlling the appearance of minor components to disclose rules for controlling major components too, and thus to increase our knowledge of the properties of common factors of traditional kinds.

GENETICAL FACTORS IN NEUROSIS*

By ELIOT SLATER

M.R.C. Psychiatric Genetics Unit, The Maudsley Hospital

My own interest in the genetical aspects of neurosis and psychopathy dates back to the last war. The circumstances then, in which a great body of men were subjected to physical and psychological stresses of a relatively stereotyped kind, to which they reacted in a variety of ways, were such as to highlight the contribution made by individual differences in personality and predisposition. Statistical investigation of a large case material (Slater, 1943) led to two conclusions.

The first of these in importance was the inverse relationship between the degree of stress undergone and the number of markers indicating abnormality of personality. The more these markers were recorded in an individual case, the less the degree of stress which had been required to produce breakdown. From this the suggestion arose that variation in neurotic predisposition was quantitative in nature, and, on the genetical side, could most plausibly be attributed to polygenes. However, the close relationship between the type of personality and the type of symptom manifested suggested the existence of qualitative as well as quantitative variation. My brother and I (Slater & Slater, 1944) subsequently formulated the hypothesis more precisely. Clinical observations could be accounted for if the disposition to neurosis was regarded as due to a number of distinct qualities, each of them subject to normal variation. An individual who differed markedly from the mean towards either extreme, in respect to any single such quality, would be more than normally susceptible to stresses of a certain specific kind; and if the stress passed a certain level of intensity he would be likely to break down, and then to show symptoms of a specific and related kind. But the majority of men breaking down under stress would, by the laws of chance, be those who showed minor deviations along a number of different lines, rather than those who showed an extreme deviation along one line. They would be liable to break down under a mixture of stresses, and to respond with a variety of symptoms. Variation in the constitutional predisposition to such reactions should correspond to a normal surface in many dimensions.

---

* A shortened form of a paper with the same title delivered at the Second International Conference of Human Genetics in Rome, 1961. Republished by permission.

A hypothesis of essentially the same kind, but much simplified, has been advanced by Eysenck, and has been supported by him and his collaborators with a great amount of valuable work. A recent formulation of his theory is given in his *Handbook of Abnormal Psychology* (Eysenck, 1960*e*). The normal surface which Eysenck uses has two dimensions, the polarities being horizontally extraversion–introversion, and vertically neuroticism–normality. With the mental tests employed, persons with anxiety states, obsessionals and depressives are found on the introverted side, psychopaths and hysterics on the extraverted side of the mid-line. On the vertical dimension all these groups cluster fairly close together, but are widely separated from the normals. A multifactorial genetical basis for the neurotic predisposition forms a natural but probably not essential part of Eysenck's theory.

In what follows, I shall endeavour to examine the question whether we can be satisfied with this theory, or whether we should prefer the concept of a pluridimensional field of variation to that of a bi-dimensional one. The question will be approached from the genetical side. If we find indications of qualitative rather than quantitative differences between groups, or if the genetical factors associated with such closely neighbouring syndromes as anxiety and obsessionality are found to be different from one another, then we shall be inclined to prefer the pluridimensional hypothesis. If no such findings are made, then the simpler view should prevail.

Of all neurotic syndromes, the evidence relating to genetical predisposition is best in the case of obsessional neuroses. Luxenburger (1930), Lewis (1935), Brown (1942) and Rüdin (1953) have all contributed to this. Luxenburger examined the parents and sibs of 71 obsessionals, and found anankastic qualities in 15 % of fathers, 6 % of mothers and 14 % of sibs. He concluded that the 'anankotrope character' was inherited as an entity. Somewhat similar conclusions were reached by Lewis. He reported on the familial data on 50 obsessional patients. Of the 100 parents 37 showed pronounced obsessional traits in one form or another, and of the 206 sibs 43, or 21 %, showed mild or severe obsessional traits. In Brown's material, nine cases of definite obsessional neurosis were found among the parents and sibs of the three neurotic groups, seven among the 96 first-degree relatives of obsessional probands, two among the 344 relatives of anxiety neurotics, and none among the 105 relatives of hysterics or the 185 relatives of control cases. Brown differed in his conclusions from other workers, since he did not think his findings supported, as in fact they do, the existence of a specific obsessional predisposition. The most recent report is that of Rüdin. She found 5 % of the parents and 2 % of the sibs of her obsessional probands had had an obsessional illness, and 5 % of the parents and 3 % of the sibs had obsessional personalities. Obsessionality in one or other form was found in 8 % of parents and 5 % of sibs. The differences between the results obtained by these workers depends in part on the criteria used.

The data relating to anxiety states are very much less in amount. McInnes (1937) examined the relatives of 50 anxiety neurotics, and found that 15 of the 100 parents and 28, or 15 %, of the 189 sibs had anxiety neuroses also. There were also small numbers of others suffering from unclassified neurotic states and other psychiatric abnormalities, but anxiety state was the only condition shown in obviously marked excess. In the relatives of psychiatrically normal control subjects, 4 % of parents and 5 % of sibs had anxiety neuroses. Brown found that 21 % of the parents and 12 % of the sibs of anxiety neurotics had also had anxiety neuroses, i.e. conditions which at some time had produced some incapacity socially or at work; but such conditions were also found with about half that frequency in the parents and sibs of hysterics and in the sibs of obsessionals. Cohen (1951) worked with army recruits suffering from 'neuro-circulatory asthenia' (NCA), i.e. a form of anxiety neurosis in which the circulatory symptoms predominated. They divided their material into 67 chronic and 44 acute cases, and compared the findings with those obtained from 18 men convalescent from physical disorders and 54 healthy soldiers. It was the group of chronic cases that showed the distinctive findings, 18 % of the fathers, 55 % of the mothers, 13 % of the brothers and 12 % of the sisters of these probands suffering also from NCA. Among the relatives of the acute cases the prevalence of NCA was hardly higher than in the relatives of the control groups. The authors considered various possible genetical models for the explanation of their findings, and thought that a dominant or double dominant hypothesis fitted the facts best. They add: 'In addition, the study of 37 children of 18 NCA parents revealed no evidence of hysteria, depression or obsessional neurosis, although there was a high prevalence of NCA in these children.'

When we turn to hysteria, we find that the evidence is conflicting. Kraulis (1931) took a group of severely abnormal individuals, who were psychopathic in a general sense as well as being subject to hysterical symptoms; 9 % of the parents and 13 % of the sibs had hysterical reactions, and there was a raised prevalence of abnormalities of personality among the sibs. Much lower figures are reported by McInnes for the families of 30 hysterical patients. Of the 60 parents two suffered from anxiety states and two from hysteria; of the 117 sibs two suffered from anxiety states and none from hysteria. Brown, on the other hand, found relatively high prevalences, 19 % of the parents and 6 % of the sibs of his hysterics having had hysterical symptoms causing social incapacity, paralyses, amnesic fugues, hysterical fits, etc. Ljungberg, in his important study (1957), reports much lower rates: in the fathers, brothers and sons of hysterics 2 %, 3 % and 5 % of hysterics; in the mothers, sisters and daughters 7 %, 6 % and 7 %. The present writer (Slater, 1961) reported 24 pairs of twins, 12 monozygotic (MZ) and 12 dizygotic (DZ), in which the proband had been diagnosed as suffering from hysteria. There was not a single co-twin who had ever been diagnosed as suffering from hysteria, though abnormalities of personality and psychiatric illness were common. Reviewing these reports, one must conclude that hysterical illness is almost certainly commoner among the first-degree relatives of hysterics than in the general population. Ljungberg thought his findings were compatible with the operation of multi-factorial genetical factors. My own conclusion was that, though this might well be true, the factors involved were probably not very specific, and that the aetiology, natural history and outcome of hysterical states were all too various for the condition to be satisfactorily regarded as a syndrome.

The search for specific factors in the functional disorders has been more successful where the symptoms could be regarded as related to a definite disturbance of normal neurophysiology. Thus, from his work on 'specific dyslexia' Hallgren (1950) concluded that genetical causation by a monohybrid dominant was probable. The same author's findings with enuresis (1957) were rather less striking, but led him to the view that there was probably a nuclear group of nocturnal enuretics in which the condition was genetically determined. In these conditions the genetical factors involved must be regarded as having some degree of specificity. This is not invariably the case. Thus Kaij (1960) has shown that MZ twins resemble one another more closely than do DZ twins in respect of drinking habits; and in Dencker's study (1958) of 128 twins with closed head injuries it was found that it was certain sorts of neurotic symptoms which were more concordant in MZ than DZ pairs, other common sequelae showing no such relationship to the inborn constitution. In these latter conditions the genetical contribution to the causation of the form of disturbance investigated could be entirely unspecific, having its main effect, say, in predisposing to emotional instability or lack of control.

Now, if there are specificities in the predisposition to neurosis, then one might expect that in pairs of twins both of whom develop a neurotic illness there would be greater qualitative similarity within MZ than within DZ pairs. My colleague, Mr James Shields, has made a number of observations bearing on this point. In his report on 36 MZ and 26 DZ pairs of normal twin schoolchildren (Shields, 1954), he shows that 42 out of the total of 124 individuals had a moderately severe degree of psychiatric maladjustment, severe enough to have at some time caused referral to some social agency. The symptoms were of the kind familiar to the child psychiatrist. Analysis of the figures, to distinguish between quantitative and qualitative concordance, showed that MZ twins were only a little more alike than DZ twins in respect of the presence or absence of neurosis or behaviour disorder, but that they were very much more alike in the type of neurotic symptom or personality trait. The occurrence or non-occurrence of neurotic disorder appeared to be largely determined by the environment; the form of the neurosis depended more on the genetical constitution.

In Shields's study (1962) on MZ twins who had been separated from one another in infancy or early childhood, both similarities and dissimilarities in neurotic symptomatology were very striking. Thus at one extreme there was a pair who had both complained at times of globus hystericus, and had been submitted to much hospital investigation for physical disorders with negative results; and at the other extreme there was a pair in which only one had been a rebellious child, had made homosexual attachments and had had three attacks of amnesia, while her twin was normal. In general, illnesses of a hysterical kind showed up in the discordant pairs, while community in respect of anxiety states accounted for several concordant pairs.

In a further study at present in progress in the M.R.C. Genetics Unit at the Maudsley, we have 14 MZ pairs and 5 DZ pairs in which both twins had some form of neurotic disorder. In six of the MZ pairs the twins showed close similarity in a number of important respects and very little difference of a qualitative kind. There were no such pairs among the DZ twins, who showed wide qualitative differences. Thus we have in the DZ pairs a chronic alcoholic paired with an exhibitionist liable to anxiety states; a chronic hypochondriac paired with a woman who had an anxiety state during a pregnancy; a chronic alcoholic and drug addict paired with an anxious hypochondriac; a chronic obsessional paired with a chronic hysteric; and a suicidally inclined homosexual man whose twin was a truant at school, later became sexually promiscuous and a shop-lifter, and had many neurotic complaints, some involving a suggestion of malingering. In the group of twins notified from the Maudsley Children's Department, whom we have been following up for a number of years, the same general observation is made: concordant MZ pairs are much more alike in their symptomatology than are concordant DZ pairs.

Reviewing the evidence, it seems to me that in the field of variation of temperament and personality we have to face much the same situation as we have in the field of variation in intellectual ability. It is a convenience for the psychologist to use the simplest possible hypothesis for the design of mental tests and the calculation of scores. By means of such simple hypotheses a useful part of the total observable variance can be accounted for. In both fields, however, as we refine our studies we find specific mental abilities and specific traits or constellations of traits of personality which find no place in such simplified models. At this point we have to broaden our working hypothesis.

PERSONAL CONTINUITY AND PSYCHOPATHOLOGICAL DISRUPTION

By G. A. FOULDS
*Runwell Hospital, Wickford, Essex**

I. THEORETICAL DISTINCTIONS BETWEEN PERSONALITY CHARACTERISTICS AND PSYCHIATRIC SYNDROMES

As Claude Bernard (1878) pointed out, the inner environment has to be more constant than the outer. In a sense in which Wordsworth probably did not intend it, the world is too much with us and we have to partial some of it out. This involves, as Bartlett (1932) showed, selection and simplification of incoming stimuli. Experience tends to get organized, consciously or unconsciously, into something which is simpler than the original. Characteristically in man, much of this organization is predigested for him by the culture in which he lives. That this selectivity serves needs deeper than man's sloth is evidenced by the fact that in the past manic patients, at the mercy of almost all stimuli, sometimes literally died of exhaustion.

---

* Now at Medical Research Council Unit for Research on the Epidemiology of Psychiatric Illness, Department of Psychological Medicine, University of Edinburgh.

Terms like 'trait' and 'attitude' were introduced into Psychology to describe certain aspects of this economy of organization which imply or lead to personal continuity. Traits and attitudes on the one hand can be distinguished from symptoms and signs of personal illness on the other by means of three criteria:

(1) Traits and attitudes are *very common*; symptoms and signs are not.

(2) Traits and attitudes are relatively *ego-syntonic*; symptoms and signs are distressful, either to the patient or to his closest associates.

(3) Traits and attitudes, particularly traits, are relatively *enduring*; symptoms and signs are relatively transient.

Thus we are all more or less quick, more or less extrapunitive and we are so, with possibly some modification in the more or less, throughout our lives and we do not particularly mind. We do not, however, all have *flexibilitas cerea* or delusions of guilt and, if we do not mind, somebody will; nor do those who do have them have them throughout their lives.

'Trait' is being used as a constellation of expressive behavioural characteristics. It does not drive. To follow McClelland's (1951) distinction, traits are concerned with the 'how' of behaviour. How did X do Y? He did it quickly, gracefully, independently. *Attitudes* are concerned with the 'what' of behaviour. What did X think or feel about Y? He felt antagonistic towards him, that he was the sort of person who would let his friend down, etc.

A *personality type* may be regarded as a constellation of strengths of traits and attitudes which can be distinguished from other such constellations. An example would be the hysteroid personality, as described by Lewis & Mapother (1941).

Personality traits, attitudes and types emphasize the continuities in behaviour; whereas the defence dynamisms, manifested in part in characteristic symptom and sign clusters, which signal a disruption of continuity, represent an attempt to reorganize the retreating forces, to maintain some sort of integration of the self concept, or self schema, and to continue to be in relation with other persons. A symptom or sign may be viewed as essentially a *change* in bodily or mental functioning.

A syndrome is a constellation or cluster of symptoms and signs which tends to occur rather commonly and which is distinguishable from other such clusters. Symptoms or signs may occur in more than one syndrome; but the syndromes *per se* must be mutually exclusive.

II. CRITERION MEASURES OF PERSONALITY CHARACTERISTICS AND PSYCHIATRIC SYNDROMES

In order to investigate this personality type–psychiatric syndrome distinction quantitatively, criterion measures had to be developed for each—these are Caine's Hysteroid-Obsessoid Questionnaire (Martin & Caine, 1963) and the Symptom–Sign Inventory (Foulds, 1962*a*).

The Hysteroid Questionnaire had a biserial correlation of 0·68 with observer ratings of the traits from which the Questionnaire was constructed and these traits were derived from the Lewis & Mapother description. When the ratings and the Questionnaire were dichotomized, 78 % of the cases were allocated in agreement. In a second study the figure was again 78 %. In the present studies the HOQ was therefore used as the criterion *personality* measure.

The criterion *diagnostic* measures were derived from the Symptom–Sign Inventory. Before going on to describe it, it will be necessary to say something about diagnostic classification.

The term 'personal illness' has been preferred to functional mental illness to indicate severe disturbance in the form of the personal rather than merely in the form of the organic. The purely personally ill may be differentiated from those who are not only personally ill and those who are not ill at all (i.e. psychosomatic, somatopsychological and organic cases and those who are altogether free from ill-health) by means of three criteria. Within the personal illness class, we are concerned with people: *first*, who are experiencing difficulty in maintaining or establishing mutual personal relationships; *secondly*, whose difficulty is so distressful to the individual that he seeks outside help in order to alleviate the symptoms which have been thrown up to defend himself against facing this basic problem (alternatively, whose symptoms are so distressful to his most intimate associates that they have sought help on his behalf for the same purpose) and, *thirdly*, whose organic concomitants, if any, of the personal illness are not such as, in the absence of the personal illness, would warrant medical or surgical interference.

The tragedy and essence of personal illness lies in the alienation from the rest of humanity. To paraphrase John Donne, no person is an island. In seeking, or being forced, so to become, the personally ill diminish themselves, or are diminished as persons. The normal integrative individual, knowing who he is, can risk revealing himself and thus allowing himself to become vulnerable and so to become capable of entering into mutual personal relationships. The disturbed, partially integrated individual, fearing the possibility of rejection, of modification, of further disintegration, cannot allow himself to become vulnerable. By protecting himself, by his inability to reveal himself, he precludes the possibility of being known by others, of knowing others and of knowing himself. He thus cannot enter into mutual personal relationships.

It may be helpful to think of a continuum stretching from normality through psychopathy, neurosis, integrated psychosis (in the main, melancholics, manics and paranoids) to non-integrated psychosis (in the main, the schizophrenias). As we proceed along this continuum there is an increasing failure in the individual to *intend* his own actions and to *share* his experience.

In terms of such a continuum, non-integrated psychosis implies psychosis and therefore personal illness, and psychosis implies personal illness. This is so logically and may possibly be so historically. This does not imply that all melancholics, for example, have obsessional symptoms. What it does imply is that melancholics have, with comparable frequency, all those symptoms which are equally common in all neurotic subgroups.

Accepting King Lear's principle that 'where the greater malady is fixed, the lesser is scarce felt', it makes sense to speak of a non-integrated psychosis—schizophrenia—with neurotic-anxiety features; but it does not make sense to speak of an anxiety state with schizophrenic features.

Table 1 shows the suggested necessary and sufficient conditions for making the major class diagnoses in terms of the type of disorder.

About 15 scales have been worked out on the Symptom–Sign Inventory; but only three will be discussed here—the *Non-integrated* v. *Integrated Psychotic* scale; the *Psychotic* v. *Neurotic* scale and the *Personal Illness* v. *Normal* scale. The results were obtained from 263 women who were psychiatrically ill and 73 who were not. The age range was 20–59.

Table 1. *The personal illness continuum, with the necessary and sufficient conditions for the major categories*

Absence of each of the following constitutes necessary conditions for mutual personal relationships	NIP	IP	N	PP
(1) *Loss of awareness of self as agent.* Delusions or hallucinations of passivity, catalepsy, etc.	S*	—	—	—
(2) *Thought process disorder.* Blocking, neologisms, knight's move, over-inclusion, word salads etc.	S*	—	—	—
(3) *Affective disorder* (difference in kind). Incongruity (including flattening)	S*	—	—	—
(4) *Perceptual disorder.* Hallucinations, gross distortions (including imperceptions)	N′	S*	—	—
(5) *Thought content disorder.* Delusions implying self-schema	N′	S*	—	—
(6) *Affective disorder.* Disproportionateness	N′	N′	N′S	—
(7) *Egocentricity.* Lack of empathy, treating people as objects	N′	N′	N′	N′S

NIP = non-integrated psychosis; IP = integrated psychosis; N = neurosis; PP = psychopathy; N′ = necessary condition; S = sufficient condition. * A necessary condition that one at least be present.

It follows from the continuum hypothesis that the three scales should be unidirectional and that, when we come to differentiate the personally ill from normals, it will be the symptoms forming the lowest common denominator, as it were, which will be the discriminators. We will expect these then to be symptoms which neurotics and psychotics have equally and more frequently than normals. Indeed, one cannot find items which are common to all types of neurotics which are not equally common in all types of psychotics.

When we come to differentiate psychotics from neurotics, the items which differentiate will be psychotic items. When we come to differentiate non-integrated from integrated psychotics, the non-integrated psychotic items will be the operative ones.

The usual practice of comparing, say, a group of neurotics with a group of psychotics has not been followed. Items would only get into this scale which differentiated (at the 5 % level at least on a χ^2 test) each of the subgroups in one class from each of the subgroups in the other class bar one, where that one was in the same direction.

There are thus 18 items which get into the Personal Illness *v.* Normal scale, of which 16 came from the *a priori* neurotic scales. There are no items given more frequently by normals than by patients.

Eighty-five per cent of the psychiatrically diagnosed non-integrated psychotics complain of four or more of these 18 items; 88 % of integrated psychotics; 89 % of neurotics and 11 % of normals.

Most of these items could be subsumed under the rubric of affective disproportion—examples at random are:

'I suffer from palpitations and breathlessness.'
'I am afraid I might be going insane.'
'I cry rather easily.'
'I am slower recently in everything I do.'
'The future seems pointless.'

On the Psychotic *v.* Neurotic scale, the percentages scoring 4 or more are: non-integrated psychotics 65; integrated psychotics 68; neurotics 16; normals 9. There is a band of scores at which it is uncertain to which class they belong.

A random sample of the 14 items is as follows:

'I have an important mission to carry out.'
'People are plotting against me through no fault of my own.'
'I sometimes hear voices without knowing where they come from.'
'I am a condemned person because of my sins.'
'I do *not* have black-outs, dizzy spells or faints.'

Twelve items were given more frequently by psychotics and only two more frequently by neurotics. Ten out of the 12 psychotic items were delusional items.

On the Non-integrated *v.* Integrated Psychotic scale, the percentages were: non-integrated psychotics 65; integrated psychotics 25; neurotics 5; normals 6. Again there was an 'uncertain' band of scores.

Random examples are:

'I am compelled to think over abstract problems again and again until I can't leave them alone.'
'I often wonder who I really am.'
'I sometimes have very strange and peculiar thoughts.'
'There is something unusual about my body, like one side being different from the other and meaning something different.'
'I have *not* been in poor physical health during the past few years.'

Thought disorder and incongruity of affect are not picked up at all well on the Inventory; but many of the items are concerned with loss of awareness of the self as agent, with loss of boundaries, of the self concept, etc.

Eleven items are given more frequently by non-integrated psychotics and only one more frequently by integrated psychotics.

Thus, non-integrated psychotics score high on all three scales; integrated psychotics on two; neurotics on one and normals on none. This lends some support to the view that non-integrated psychosis implies psychosis, which implies personal illness, and that psychosis implies personal illness. That this is not just an artefact of the inventory is suggested by the fact that agreement with the psychiatric diagnosis on the three scales is high.

III. EMPIRICAL DISTINCTIONS BETWEEN PERSONALITY CHARACTERISTICS AND PSYCHIATRIC SYNDROMES

Using these scales derived from the Inventory as criterion diagnostic measures and the Hysteroid-Obsessoid Questionnaire as the criterion personality measure, a study will be described in which two other measures are related to the two criterion measures. The additional measures are *General Punitiveness* and *Direction of Punitiveness* (whether a person tends to be predominantly extra- or intro-punitive).

It is predicted that the measure which is significantly associated with the criterion diagnostic measure will change on re-test, at least in those patients rated by psychiatrists as much improved after 4–6 weeks, whereas the criterion personality measure and the measure significantly associated with it and not with diagnosis will not change. This, of course, follows from the earlier assertion that personality traits and attitudes are relatively more enduring than symptoms and signs.

The subjects were all women, aged 20–59. There were 20 cases in each of the following groups (psychiatric diagnosis being used for the purpose): schizophrenics (non-paranoid); paranoid states (including paranoid schizophrenics); melancholics; hysterics; anxiety states and neurotic depressives. All 120 cases were new, but not necessarily first admissions. All 120 were tested within one week of admission and 88 were re-tested after 4–6 weeks.

Table 2 shows the correlations between the three diagnostic scales Non-integrated Psychosis (NIP) *v.* Integrated Psychosis (IP); Psychosis (P) *v.* Neurosis (N) and Personal Illness (PI) *v.* Normal (No); the Hysteroid-Obsessoid Questionnaire (HOQ); General Punitiveness (E+I) and Direction of Punitiveness (E−I).

Table 2. *Correlation between diagnostic and personality criterion measures and measures of punitiveness*

	(E+I)	(E−I)	HOQ
NIP *v.* IP	0·396*	0·025	0·039
P *v.* N	0·532*	0·134	−0·015
PI *v.* No	0·392*	−0·187	−0·147
HOQ	−0·017	0·408*	—

* $P < 0{\cdot}001$.

Thus, non-integrated psychotics are more generally punitive than integrated psychotics, who, in turn, are more generally punitive than neurotics—which fits in with what was said about a continuum of increasing failure to maintain mutual personal relationships. None of these diagnostic scales is significantly related to direction of punitiveness. On the other hand, the criterion personality measure—the HOQ—is significantly correlated with direction of punitiveness—the more hysteroid being more extrapunitive.

In view of the emphasis on failure in mutual personal relationships, the general punitive measure is perhaps of particular interest. It may be a reasonable estimate of such failure. The more successful people are in their personal relationships the less they resort, even at times of stress, to blaming either themselves or others. As Anderson (1939) has pointed out, such people seek an integrative solution. This fact emerges too from the therapeutic work of Rogers and his colleagues.

From the argument about the continuum of failure in personal relationships, the expected general punitive order would be normals, neurotics, integrated psychotics, non-integrated psychotics. Table 3 shows the means derived from a combination of two studies.

Table 3. *Mean general punitive scores* ($E+I$) *of three major psychiatric and one normal group of women*

20 non-integrated psychotics	25·75
80 integrated psychotics	20·04
120 neurotics	17·37
32 normals	11·72

Table 4. *Test score changes in relation to clinical improvement*

	37 improved				51 slightly or not improved			
	Mean	S.D.	*t*	*P*	Mean	S.D.	*t*	*P*
SSI	12·65	10·86	7·08	<0·001	4·88	7·85	4·44	<0·001
E+I	6·89	8·32	5·03	<0·001	0·82	5·71	1·03	N.S.
E−I	−1·00	4·47	1·36	N.S.	0·18	3·88	0·33	N.S.
HOQ	−1·38	4·13	2·03	N.S.	0·37	4·60	0·57	N.S.

We can say now that symptoms and general punitiveness should drop significantly, at least in those rated as much improved. Direction of punitiveness and HOQ should *not* change significantly.

Table 4 shows the mean difference scores for test and re-test.

The total number of symptoms dropped significantly in both groups, but less in the slightly improved or unimproved group. In all other respects the measures behaved according to prediction.

Summarizing: (1) personality traits and attitudes are very common, relatively ego-syntonic and relatively enduring in contrast to symptoms and signs of personal illness; (2) it may be useful to think of a continuum of increasing failure to maintain or establish mutual personal relationships through psychopathy, neurosis, integrated psychosis to non-integrated psychosis. In support of this, non-integrated psychotics, as a group, have all the symptoms (on the SSI) that are common to the other groups plus their own; integrated psychotics have the symptoms that are common to all neurotics plus those shared with non-integrated psychotics and neurotics have symptoms in common with the psychotics, which differentiate them from normals; (3) general punitiveness correlates with the diagnostic criterion measure, but not with the personality criterion measure, and this measure changes significantly after 4–6 weeks in hospital; (4) direction of punitiveness correlates with the personality criterion measure, but not with the diagnostic criterion measure and this measure does not change significantly after 4–6 weeks.

SOME POSSIBLE IMPLICATIONS

Any classificatory scheme is only one out of many possible consistent orderings of complex data to be judged by their pragmatic value. It has been suggested that it may be useful to regard: (*a*) psychiatric syndromes as falling along a continuum of increasing failure to maintain mutual personal relationships; (*b*) symptoms as distinguishable from personality traits and attitudes.

Acceptance of this position might lead to:

(1) A decrease in diagnostic confusion. From (*a*) would follow, for example, recognition that the diagnostic category should be based on the symptoms falling at the position farthest removed from the normal end of the continuum. From (*b*) would follow, for example, rejection of the practice of diagnosing cases, along with, say, functional paralyses and fugues, as hysteria who merely exhibit lability of affect, histrionic behaviour and a tendency to manipulate other people.

(2) A more refined assessment of treatment effects. Thus, some psychotics, for example, are discharged when still psychotic; some when still neurotic, but no longer psychotic; some when basic personality attitudes have and have not been modified. Follow-up studies of these different groups might be illuminating.

(3) A clearer basis for genetic studies. Thus, the presence of obsessoid personality traits in one monozygote should not be taken as evidence of concordance with an obsessional neurotic twin of hysteroid personality.

(4) In general, a greater clarification of what precedes breakdown, what survives during breakdown and what remains after remission.

PERSONALITY IN THE DIAGNOSIS OF HYSTERIA

BY J. G. INGHAM AND J. O. ROBINSON
Medical Research Council, Social Psychiatry Research Unit,
Detachment at Llandough Hospital, Penarth, Glamorgan

INTRODUCTION

This paper is concerned with the hypothesis, suggested by Jung, that hysteria is the typical neurosis of extraverts and that psychasthenia is the typical neurosis of introverts. Eysenck equates his concept of dysthymia with the earlier one of psychasthenia. In his terminology the hypothesis is that individuals with extraverted patterns of behaviour tend to develop hysteric-psychopathic disorders in cases of neurotic breakdown and that individuals with introverted patterns of behaviour tend to develop dysthymic disorders in cases of neurotic breakdown.

Hysteric-psychopathic and dysthymic disorders are defined in terms of psychiatric observations. For the moment, let us assume that a hysteric is a patient who by a consensus of psychiatric opinion would be labelled as such and a dysthymic is a patient who, by a consensus of psychiatric opinion,would be labelled as suffering from anxiety state, neurotic depression, obsessive-compulsive disorder, or phobia.

The definition of extraverted and introverted patterns of behaviour has been the topic of some recent controversy (Foulds, 1961, 1962; Eysenck, 1962; Ingham, 1962). Foulds (1961) expressed the view that it was not legitimate to use hysterics and dysthymics as criterion groups to test the validity of an extraversion scale. But, for the purpose of this discussion, it is assumed that extraversion is the variable so named in questionnaire studies by Hildebrand (1958), Eysenck (1956*b*) and others. This has the great advantage of being observable by precisely defined methods. The Maudsley

Personality Inventory (M.P.I.) was the questionnaire used in the present work and extraversion is defined as the E-score of the M.P.I. The corresponding neuroticism (N) scores are also available.

PREVIOUS WORK

Before describing the results, let us consider some previous investigations. They are summarized in Fig. 1. The horizontal axis represents the E-score; the vertical axis, the N-score. These axes have been drawn at right-angles for clarity but this should not be taken to imply that the two scores are uncorrelated. It is now well established that, in a population of neurotic patients in hospital, there is a negative correlation (something like $-0{\cdot}4$) between them. Each letter shows the mean scores of one sample, from one investigation.

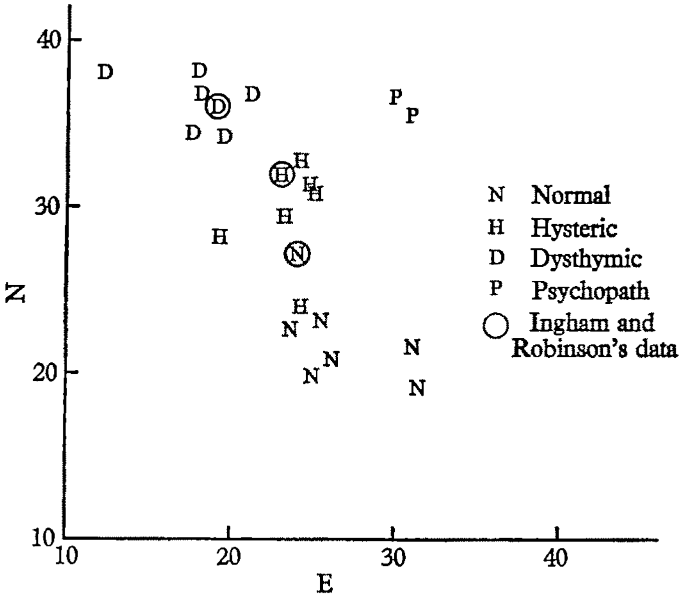


Fig. 1. Mean M.P.I. N- and E-scores observed in eight investigations including the present one. Based upon results published by Claridge (1960), Claridge (1961), Claridge & Herrington (1960), Eysenck (1959), Eysenck (1961), McGuire, Mowbray & Vallance (1963), and Sigal *et al.* (1958).

These results, with one exception, are in line with predictions made from the hypothesis. Dysthymics tend to be introverted compared with hysterics, psychopaths or normals. The normal samples, however, seem to be as extraverted as hysterics. This is not in line with Eysenck's original prediction and has evoked a certain amount of discussion (e.g. Sigal, Star & Franks, 1958). As far as neuroticism is concerned, all the neurotic groups have higher N-scores than the normal groups. It is also worth noting that the hysterics seem to be rather less neurotic than either dysthymics or psychopaths.

The scores obtained in the present work also appear in Fig. 1. The hysteric and dysthymic groups had mean scores which came well within the range of previous investigations. There were not enough psychopaths to form a separate group. The normal group differs markedly, however, from those of previous investigations in having a considerably higher mean N-score. There are several possible explanations for this. One is that there are real differences between the different populations from which the subjects were drawn. If this were so the neurotic groups should also differ from those of previous investigations. Since they do not a more plausible explanation may be found in differences in sampling techniques. The normal group was a random sample of the adult population of a South Wales mining valley (obtained through the courtesy of the M.R.C. Epidemiology Research Unit). In most previous investigations the normal samples have comprised either selected occupational groups or quota samples selected for market research. Possibly the nearest approach to a random sample was that described by Eysenck (1961). In this instance, 944 subjects were approached by post. They were selected at random from the members of a panel used by a large market research organization 'with substitution for refusals'. The original refusal rate is not stated, but with this method it was probably fairly high. If the people who refused contained a higher proportion of neurotics, this could account for its having lower mean N-scores than our sample, in which the refusal rate was low. It is true that in Eysenck's investigation there was a follow-up several months later and the original scores of those who responded on the second occasion were compared with those who refused on the second occasion. There was no difference between these two groups, but this does not establish that those who refused initially were not more neurotic. Another possible explanation of the high N-scores is that a personal administration of the inventory, in a face-to-face interview situation without the help of other members of the family, may well produce different scores compared with those obtained by postal contact.

Comparisons of different diagnostic groups rely heavily on psychiatric diagnosis. But there is a great deal of variation between psychiatrists on the question of diagnosis and it would be useful if one could specify precisely how a patient has been placed in a particular diagnostic group. Two things would be involved in such a specification. First, one would need to know what information the diagnostic decision was based upon. Secondly, one would need to know how the information was used in arriving at the decision. In science, if one has a precise idea of the information required it should be possible to specify the methods of observation sufficiently closely to enable any scientist with the necessary skill and apparatus to make the same observations. The method of observation in this case is the psychiatric interview, and there is no reason why this should not follow a specified procedure. In practice, however, reports of investigations comparing psychiatric groups rarely do specify the interview procedure used in making the diagnoses. But even if it were possible to specify a standard interview procedure acceptable to all psychiatrists, it is unlikely that any two would agree completely on the steps to be taken in passing from the information to the diagnostic decision. It is doubtful whether any one psychiatrist could specify exactly what processes had led him from observations to decision. And of course between psychiatrists there are substantial areas of disagreement.

This is apparent in the definition of hysteria. A diagnosis of hysteria is based among other things on the presence of 'conversion' symptoms and on the presence of a 'hysterical' personality or hysterical attitude towards the illness. One source of variation between psychiatrists is in the relative weighting of these categories of information. In particular there is a great deal of variation in the extent to which personality is taken into account. There are psychiatrists who only make a diagnosis of hysteria when certain classical symptoms (such as paralysis) are present and there are others who call all histrionic and egocentric patients hysterics irrespective of symptomatology. Some distinguish between conversion hysterics, diagnosed primarily on the basis of physically manifested symptoms, and hysterical personalities, diagnosed primarily on the basis of observations of general behaviour patterns. It might be argued that, if this distinction is made clearly and consistently, then the definitions are sufficient. But the observations on which the diagnoses are based are not independent. For example the likelihood of a given symptom being called a conversion symptom may, for some psychiatrists, be greater if it occurs in a patient with a hysterical personality. This interaction between symptom judgements and personality judgements will vary from one psychiatrist to another. In fact one view is that almost any symptom, physical or psychological, may be hysterical. For a psychiatrist holding this view, the diagnostic decision will be based largely on personality observations. He might still use the term conversion hysteria but not include in the group patients who have only symptoms and no indications of hysterical personality. These considerations suggest that in the investigation of differences between diagnostic groupings one ought to look closely at diagnostic criteria. What follows is an attempt to do so.

PRESENT STUDY

Subjects

The results were obtained in the course of several investigations, primarily concerned with other problems, from a group of normal subjects and several groups of neurotic patients as follows:

(1) A random sample obtained from the population of the Rhondda Fach, a South Wales mining valley. There were 78 males and 64 females between the ages of 14 and 65. These people were not involved in the psychiatric part of the investigation.

(2) Neurotic outpatients from the Psychiatric Clinic at the East Glamorgan Hospital. There were 25 male and 31 female patients all living in the Rhondda Fach or other similar mining communities.

(3) Two separate groups of neurotic inpatients, including some day hospital patients, from Whitchurch Hospital, Cardiff. There was a total of 109 males and 84 females in the two groups. They comprised all the neurotic patients coming to the hospital during the periods covered by the two investigations concerned. Most of these patients did not live in a mining valley and in consequence probably differed from the others in some respects. The mean M.P.I. scores of the Rhondda patients were remarkably close to those of the Cardiff patients, however, and it seems justifiable to combine them.

Interview

All the patients were interviewed by one of us. The aim of the interview was to get enough information for the interviewer to complete a proforma like that shown in the Appendix. The interview was divided into two phases. During the first the patient was asked to give an account of his illness by a question such as 'What is it all about?' 'What's been troubling you?' or 'What's the trouble exactly?' He was then allowed to tell his story with as little intervention as possible from the interviewer. Questions were asked with the sole object of guiding the patient to those aspects of life on which information was required. Particular care was taken to avoid any direct questions relating to items in the proforma. For example, the question 'How did you get on with your parents?' was permissible but 'Do you feel anxious?' was not.

The following areas were covered in the interview:

(1) The immediate illness; symptoms and troubles associated with it; the circumstances surrounding its onset; its course to date.

(2) Occupational history.

(3) Childhood: schooling, relationships with parents, siblings and friends.

(4) Social relationships.

(5) Marital relationship; relationships with own children; sexual history.

(6) Feelings about the future.

The patient's general behaviour was observed throughout, particularly for signs of anxiety such as strained expression, fidgeting and trembling. As much information as possible was recorded in note form.

This stage usually took from 30 to 45 minutes but the patients were allowed to go on talking until it was felt that no further useful information was appearing. The in-patients were then asked to do the Maudsley Personality Inventory and, while they were doing this, the proforma was completed, as far as it could be, using the information already obtained. The out-patient group, like the random sample, did the M.P.I. as one of a series of tests on a separate occasion. When the patients had finished the M.P.I. a list of questions, referring specifically to proforma items, was asked. Answers were recorded verbatim. For example, one of the questions was 'Do you sometimes find your heart beating very quickly or very hard?' For certain proforma items, therefore, there were two kinds of information, one obtained in the first phase of the interview without specific questioning and the other contained in response to questions. The two kinds of information were recorded separately so that it was known, for example, whether the patient had complained spontaneously of depression or whether he answered 'Yes' to the question 'Do you sometimes feel depressed?' An effort was made to avoid thinking of the diagnosis until all the other questions had been answered, but it would be difficult to say how far this was achieved. The diagnosis was made last and was made with reference to all the information obtained during the interview, including that summarized in the proforma (see Appendix). It was deliberately a 'clinical' diagnosis; i.e. there was no attempt to introduce a mechanical combination of the proforma categories. The following diagnostic definitions were used for hysterical disorders (from Shagass & Naiman, 1956).

Conversion hysteria. Only patients who presented with major somato-sensory conversion symptoms. These include motor paralyses, amnesia, blindness, deafness, loss of consciousness, paraesthesia and globus.

Hysterical personality. These patients did not present with major conversion symptoms, the most common complaint being either anxiety or depression. However, during the examination and subsequent stay in hospital, each showed marked histrionic and manipulative behaviour and gave clear evidence that presenting symptoms were of a communicative nature.

Similar brief definitions of anxiety state, neurotic depression, obsessive-compulsive neurosis and phobia were used. All these were included in the dysthymic group of disorders. There were also patients with mixed neurosis having features of anxiety, depression and hysteria without very clear predominance of any one of them. The process of diagnosis consisted in matching each case against the set of brief definitions and deciding which category was most appropriate. This was done as soon as possible after the interview, using all the information available. We believe that this is a similar process to that used by the clinician when he makes a diagnosis.

RESULTS

Fig. 2 shows the hysterics, separated into conversion hysterics and hysterical personalities, compared with the other groups. The dysthymic group contained only anxiety states and neurotic depressions because there were too few obsessive-compulsive neuroses and phobias. The mixed neuroses were not included. There were 26 conversion hysterics, 30 hysterical personalities and 105 dysthymics. Analysis of variance showed that there was a significant difference in E-score between normals and dysthymics and between hysterical personalities and dysthymics. On N-scores, each of the upper groups on Fig. 2 differed significantly from each of the lower groups. In other words, when all hysterics were grouped together, as in Fig. 1, they appeared to behave very much like the hysterics in other investigations; when they were considered as separate groups they seemed to behave differently. Of the two hysteric groups it was the hysterical personalities who most nearly fulfilled Eysenck's predictions. They appeared to be more neurotic than the normals on the M.P.I. but, as regards extraversion, there is no evidence that they differed from normal individuals. Conversion hysterics did not appear to differ significantly from normal individuals in either score. There is a slight suggestion that they were introverted compared with normals but this difference was not significant.

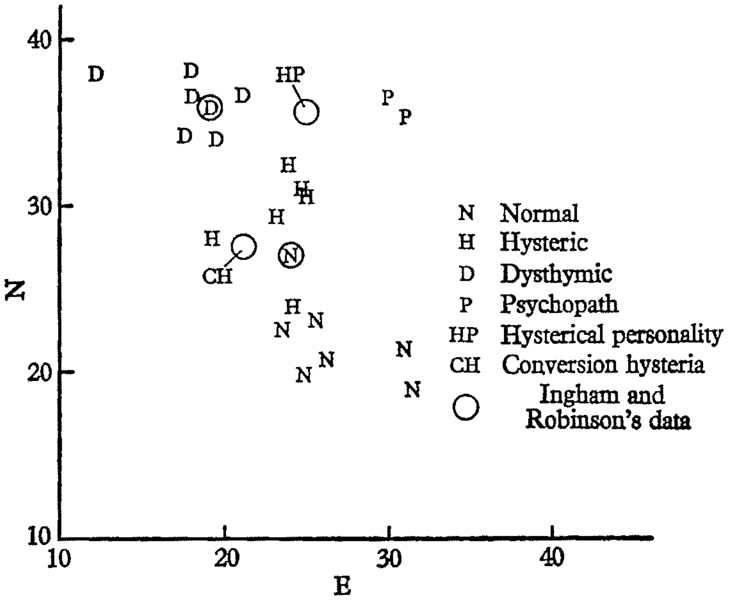


Fig. 2. The same results as those in Fig. 1 with the hysterics of the present investigation subdivided into conversion hysterics and hysterical personalities.

It seems that comparisons between the M.P.I. scores of different diagnostic groups may be very sensitive to varying diagnostic criteria so it may be a good idea to examine these criteria more closely to see what we were doing when we classified the patients. Table 1 shows contingency coefficients for the association between each interview index and the diagnosis of hysteria or dysthymia on the basis of the interview.* For example, a patient judged at interview to have a hysterical personality is significantly more likely to be diagnosed as a hysteric than as a dysthymic. This could be interpreted as an indication of the extent to which that particular item of information was taken into account in making the diagnosis.

It is worth considering the extent to which each of these indices was associated with extraversion. Table 2 shows those interview indices which were significantly associated with diagnosis in Table 1. This shows that patients with hysterical personalities (interview assessment; not diagnosis) and also those with hysterical attitudes towards their illnesses tended to have higher E-scores than other patients. It also shows that patients with skeletal indications of anxiety (i.e. trembling, etc.), those who complained of depression, those who complained of anxiety and those with anxious personalities tended to have lower E-scores than the remainder. All of these differences were statistically significant. Patients with classical hysterical symptoms, either those mentioned spontaneously or those elicited by questioning, had slightly lower E-scores than other patients but these differences were not statistically significant.

---

* Tables 1–3 involve only the last of the three groups of patients. The method of recording interview information had been improved and the results could not be conveniently combined with the earlier ones. The M.P.I. scores of all three groups were, however, very similar.

Table 1. *Contingency coefficients between interview indices and diagnoses*

	Male (83)	Female (63)
Anxious personality	0·26*	0·23
Hysterical personality	0·37*	0·35*
Anxiety feelings	0·26*	0·21
Anxiety feelings (question)	0·09	0·13
Depression	0·18	0·31*
Depression (question)	0·11	0·08
Autonomic anxiety	0·13	0·02
Autonomic anxiety (question)	0·03	0·00
Skeletal anxiety	0·21*	0·08
Skeletal anxiety (question)	0·07	0·06
Gastro-intestinal symptoms	0·12	0·03
Gastro-intestinal symptoms (question)	0·03	0·20
Classical hysterical symptoms	0·37*	0·38*
Classical hysterical symptoms (question)	0·35*	0·09
Other physical symptoms	0·18	0·06
Other physical symptoms (question)	0·11	0·14
Compulsive symptoms	0·04	0·00
Compulsive symptoms (question)	0·04	0·09
Anxious attitude	0·00	0·18
Hysterical attitude	0·30*	0·44*
Precipitating factors	0·00	0·00
Physical illness	0·00	0·00

* Significant at 5 % level.

Table 2. *Relationship of interview information with E-score and diagnosis*

	Association with E-score	P	Probability of hysteria greater when index is	Probability is increased most for
Hysterical attitude	+	$< 0·001$	+	HP
Classical hysterical symptoms (questions)	−	N.S.	+	CH
Classical hysterical symptoms	−	N.S.	+	CH
Skeletal anxiety	−	$< 0·05$	−	=
Depression	−	$< 0·01$	−	HP
Anxiety feelings	−	$< 0·05$	−	CH
Hysterical personality	+	$< 0·05$	+	HP
Anxious personality	−	$< 0·001$	−	HP

Table 3 shows the number of patients in each category on whom a positive or negative judgement was made for one particular interview item, hysterical attitude. The proportions are shown in brackets; 0·04, for example, is 5 divided by 117. Looking first at the hysterical personality category: there was a higher proportion of patients diagnosed as hysterical personalities amongst those who had a hysterical attitude than amongst those who had not. The same is also true for conversion hysterics but the difference in proportions is greater in the former case. In other words, a hysterical attitude increased the probability of a diagnosis of hysterical personality more than it increased the probability of a diagnosis of conversion hysteria. To indicate the fact, HP has been placed in the last column of Table 2, alongside hysterical attitude.

Looking at Table 2 again, all the indices shown here influenced the probability of a patient's being diagnosed hysteric. The last column shows which of the two hysteria diagnoses was influenced most using the criterion of which Table 3 is an example. Most of the attributes significantly associated with extraversion tended to increase the probability of a diagnosis of hysterical personality more than that of conversion hysteria. The presence of classical hysterical symptoms (both questions and spontaneous remarks) showed no association with extraversion and, of course, tended to increase most the probability of a diagnosis of conversion hysteria.

Table 3. *Distribution of hysterical attitude among diagnostic categories*

		Diagnostic category			
		Dysthymics +mixed	Hysterical personalities	Conversion hysterics	Total
Hysterical attitude	−	99 (0·85)	5 (0·04)	13 (0·11)	117
	+	14 (0·48)	9 (0·31)	6 (0·21)	29
Total		113 (0·77)	14 (0·10)	19 (0·13)	146

CONCLUSIONS

To summarize these results, it is clear that the evidence does not support the grouping of all hysterics into a single category. It seems more likely that they should be regarded rather as two groups, one based primarily on personality observations and the other based primarily on symptomatology. The conversion hysterics do not seem to differ very much from normal individuals with respect to M.P.I. scores. One possibility is that these patients have successfully 'converted' their neuroticism (i.e. the emotional distress which tends to produce high N-scores) into physical symptoms.

As far as E-scores are concerned, the evidence suggests that, within a neurotic population, low scores are associated with anxiety and depression and high scores with a hysterical personality and hysterical attitude. Neurotics as a whole tend to have a lower mean E-score than normal individuals and the relatively high-scoring patients with hysterical personalities or hysterical attitudes tend to have about the same mean scores as normals. Finally, there is no evidence that the occurrence of classical hysterical symptoms is associated with extraversion.

We are indebted to the M. R. C. Epidemiology Research Unit and to the East Glamorgan and Whitchurch Hospitals for enabling us to obtain the subjects for this investigation. The work was started while the writers were on the staff of the M.R.C. Neuropsychiatric Research Unit.

APPENDIX

Name Age No.

Diagnosis

Personality. In particular is patient an hysterical (e.g. egocentric, histrionic, emotionally shallow, etc.), and/or anxious, and/or obsessional type?

What does he/she complain of?

Does patient experience feelings of anxiety, insecurity, etc., to a pathological degree?

Does patient experience feelings of depression to a pathological degree?

Any autonomic manifestations associated with abnormal anxiety; tachycardia, sweating, vasomotor lability, etc.?

Any skeletal signs of anxiety, tremor, strained expression, fidgeting, etc.?

Any psychosomatic symptoms; dyspepsia, constipation, diarrhoea, amenorrhoea, etc.?

Any classical hysterical symptoms; paralyses, anaesthesias, fugues, amnesia, globus, etc.?

Any other physical symptoms?

Any compulsive symptoms?

Attitude towards illness or symptoms; 'belle indifférence', dramatization, exaggeration, apprehension, fear, etc.?

Any precipitating factors?

Any motivating factors?

Any history of relevant physical illness or injury?

Any evidence of relevant physical illness or injury now?

PRINCIPLES AND METHODS OF PERSONALITY DESCRIPTION, CLASSIFICATION AND DIAGNOSIS

BY H. J. EYSENCK

Institute of Psychiatry, University of London

Classification is an absolutely fundamental part of the scientific study of human personality; a satisfactory typology is as necessary in psychology as was Mendeleev's table of the elements in physics (Eysenck, 1954*c*). This has, of course, always been recognized, and almost everyone is acquainted with the famous typological classification into melancholics, cholerics, sanguines and phlegmatics dating back to Galen and even earlier. As this system still has much to teach us, I shall present it as Fig. 1; the outer ring in this figure shows the results of a large number of factor analytic studies of questionnaires and ratings (Eysenck, 1960*a*). As is customary in these diagrams, the correlation between any two traits is equal to their scalar product, that is to say, in this case, the cosine of their angle of separation.

Fig. 1 immediately confronts us with some of the main problems of classification. The first of these may be phrased in terms of the question: 'Categorical or dimensional?' Kant, to whom this system owes much of its popularity during the last two hundred years, was quite specific in maintaining the categorical point of view, i.e. the notion that every person could be assigned to a particular category; he was a melancholic, or a phlegmatic, or a sanguine or a choleric, but any mixtures or admixtures were inadmissible. This notion of categories is, of course, similar to the psychiatric notion of disease entities and their corresponding diagnoses; hysteria, anxiety state, paranoia, obsessional illness, and so on, are often treated as categorical entities in this sense.

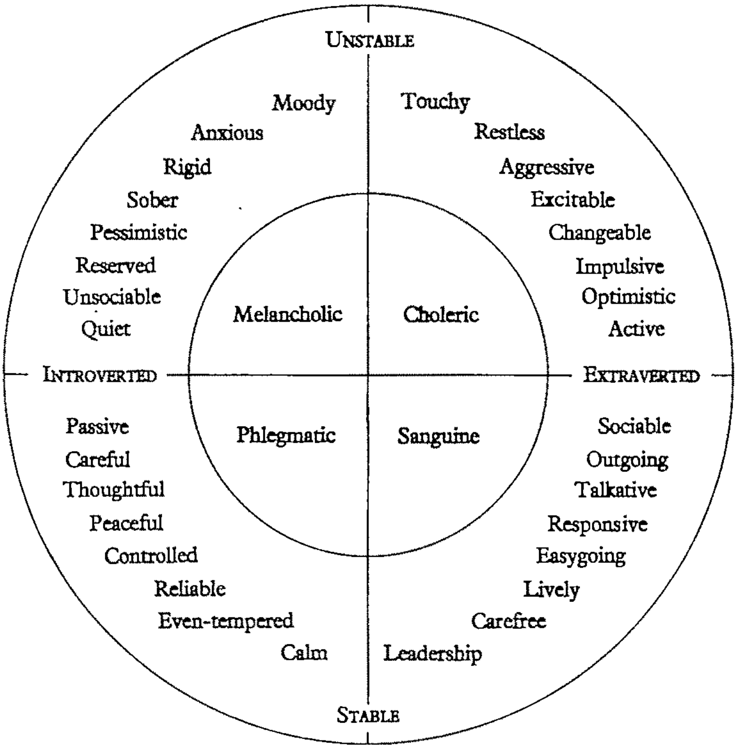


Fig. 1. Diagram showing relation between the classical four temperaments and results of modern factor analytic methods of personality description.

Opposed to this notion we have the view that any particular position in this two-dimensional framework is due to a combination of quantitative variations along the two continua labelled 'introversion–extraversion' and 'stable–unstable'. Wundt (1903), who is the most notable proponent of Galen's system in modern times, favours the dimensional view; he labelled the one axis 'slow–quick' instead of 'introversion–extraversion', and the other 'strong–weak' instead of unstable and stable.

It may be interesting to quote Wundt's very modern-sounding discussion:

> The ancient differentiation into four temperaments...arose from acute psychological observations of individual differences between people.... The fourfold division can be justified if we agree to postulate two principles in the individual reactivity of the affects: one of these refers to the *strength*, the other to the *speed of change* of a person's feelings. Cholerics and melancholics are inclined to strong affects, while sanguinics and phlegmatics are characterized by weak ones. A high rate of change is found in sanguinics and cholerics, a slow rate in melancholics and phlegmatics.
>
> It is well known that the strong temperaments...are predestined towards the *Unluststimmungen*, while the weak ones show a happier ability to enjoy life.... The two quickly changeable temperaments...are more susceptible to the impressions of the present; their mobility makes them respond to each new idea. The two slower temperaments, on the other hand, are more concerned with the future; failing to respond to each chance impression, they take time to pursue their own ideas. (Pp. 637, 638.)

There is no reason to believe that the notion of a *typology* presupposes a categorical system; both Jung and Kretschmer, who were probably the best known typologists of the inter-war period, postulated a dimensional rather than the categorical system. The widespread notion that typologists imply discontinuities, bimodal distributions, and the like, does not accurately represent the writings and views of modern typologists.

Most writers on the subject of personality come down in favour of either the categorical or the dimensional point of view without basing themselves on any experimental demonstration. I have always felt that this is unwise and that it should not be impossible to devise experimental and statistical means for verifying the one and falsifying the other hypothesis. I have tried to do this in terms of the method of *criterion analysis*, which relies on separate factor analyses of intercorrelations between tests administered to two or more criterion groups (say normals and psychotics), and the comparison of the factors emerging with a criterion column derived by serial correlation between the tests and the criterion (Eysenck, 1950). The results of this method have in every instance supported the doctrine of continuity, and failed to support the doctrine of categorization, even when the latter seemed most firmly entrenched, as in the case of psychosis (Eysenck, 1952*b*).

Assuming for the moment, therefore, the doctrine of dimensionality, we are required to build up on an experimental and statistical basis a quantitative system of personality description (Eysenck, Eysenck & Claridge, 1960). The most widely used tool for this purpose is, of course, factor analysis, and the main results of the application of this tool are shown in Fig. 1. It is notable that for many years factor analysis has been criticized because, so it was said, there was no agreement between factor analysts. Whatever may have been true twenty or thirty years ago, there can be no doubt that nowadays there is comparatively little disagreement between investigators in this field. Cattell's most recent book (Cattell & Scheier, 1961) shows him in firm agreement with the system I first put forward in 1947 (Eysenck, 1947), and Guilford, too, now appears to recognize the existence of these two main factors in personality description which I have used as the major axes in Fig. 1. Vernon (1953, p. 13) also puts forward a similar scheme. Equally we are all agreed that each of these factors is what Thurstone called a 'second-order factor', i.e. is extracted from the intercorrelations between 'first-order factors' or traits. It is with respect to these traits that much research is still needed before any final agreement is reached. Nevertheless, the major outlines of the picture are certainly beginning to appear, and it is notable that this agreement has been reached between workers using different premises, different factor analytic methods, different subjects, different tests and questionnaires, and different methods of rotation.

If we accept the principle of continuity, then we should be able to find a place for the major psychiatric classification of neurotic disorders within our Fig. 1. The theory has been put forward that neurotics suffering from anxiety, reactive depression, obsessions, phobias, and so on, would be found in the 'melancholic' quadrant, while hysterics and psychopaths would be found in the 'choleric' quadrant; psychotics would lie on an axis orthogonal to both E and N (Eysenck, 1952*a*). Descriptively there seems little doubt about the truth of this hypothesis at least as regards the neurotic groups; it is only necessary to look at the traits characterizing people in these two quadrants to realize that they might almost have been quoted from a psychiatric text-book, rather than being the result of factor analytic studies of normal people. Nevertheless, more experimental support would seem to be required. Such support, in so far as it is based purely on descriptive measures, does not remove us from some of the difficulties implied in the use of the factor analytic method. It has often been shown, as for instance in the literature deriving from the Maudsley Personality Inventory (Knapp, 1962), that hysterics, psychopaths and various dysthymic groups are in fact all high on neuroticism or emotionality, but are differentiated very significantly with respect to extraversion and introversion.

However, on a more fundamental level we may still be bothered by what is in fact the second major problem posed by our Fig. 1. This problem relates to the exact position of the axes. Mathematicians and statisticians would agree that it is perfectly legitimate to use scalar products to indicate the relative position of two traits in the dimensional space indicated in Fig. 1, and they would also agree that the position of the traits can be legitimately referred to any two arbitrary axes drawn at right-angles in the plane. They would not, however, agree with the claim sometimes made that the position of these axes can be determined in any but an arbitrary or trivial sense by statistical or mathematical considerations *alone*, as is suggested by many psychologists, particularly in the United States. I have always agreed with this criticism and have tried to argue that by retaining purely statistical criteria of axes psychologists have got themselves separated off from the main body of experimental psychology, and have remained cocooned within a small tail-chasing system incapable of generating hypotheses that could be falsified (Eysenck, 1952*c*, 1953, 1954*a*, *b*, 1956*a*). What then is the answer to this problem?

My suggestion would be that a purely descriptive system in science inevitably must carry the burden of subjectivity, and that it is because they have only been interested in *description* that factor analysts have failed to make a major impact on psychology. What is required, so I would maintain, is a set of theories linking the major aspects of the descriptive system to *causal* theories which would be capable of falsification (Eysenck, 1957). As an example of what I have in mind I may perhaps mention the set of theories relating introversion to heightened cortical excitation and lowered cortical inhibition, and extraversion with lowered cortical excitation and heightened cortical inhibition. This enables us to make large numbers of predictions of an experimental nature which are unlikely to be verified unless both the descriptive and the causal systems, and the relations specified to exist between them, are in fact in some degree related to reality. Many such predictions have in fact been made, and the great majority have been verified; I may refer in this connexion to hypotheses such as that extraverts, as compared with introverts, would be more difficult to condition, have larger reminiscence scores, have greater pain tolerance but less tolerance for sensory deprivation, are more subject to satiation, have lower sedation thresholds, have greater alpha frequency and amplitude on the EEG, more involuntary rest pauses during massed practice, have poorer vigilance, have greater speed/accuracy ratios, shorter after-images, and so forth.

This differentiation between *descriptive* and *causal* is of course related to that between *phenotypic* and *genotypic*, first made in the personality field by Pavlov on the basis of some of his animal experiments. (For a discussion of Pavlov's views and their development, Teplov's very interesting account may be consulted with advantage;

it is available in English together with a detailed evaluation of recent Russian work in the personality field (in Gray, 1964).) I have tried to indicate the difference, from the point of view of personality structure, in Fig. 2; a detailed discussion of this point is given elsewhere (Eysenck, 1960*b*). In this diagram, the subscripts '*C*' and '*B*' refer to *constitution* and *behaviour* respectively; '*E*' refers to environmental influences. It will be seen that at the most fundamental level we have the constitutional concept of the excitation/inhibition balance, which may be tilted in one direction or the other to give rise to constitutional, genotypic differences in extraversion–introversion; these may with some degree of accuracy be measured in terms of conditionability, vigilance, figural after-effects, and other laboratory phenomena. Observable behaviour is a function of these constitutional differences in interaction with the environment; this interaction gives rise to descriptive, phenotypic differences in extraversion–introversion, which can best be measured in terms of questionnaires such as the M.P.I.*

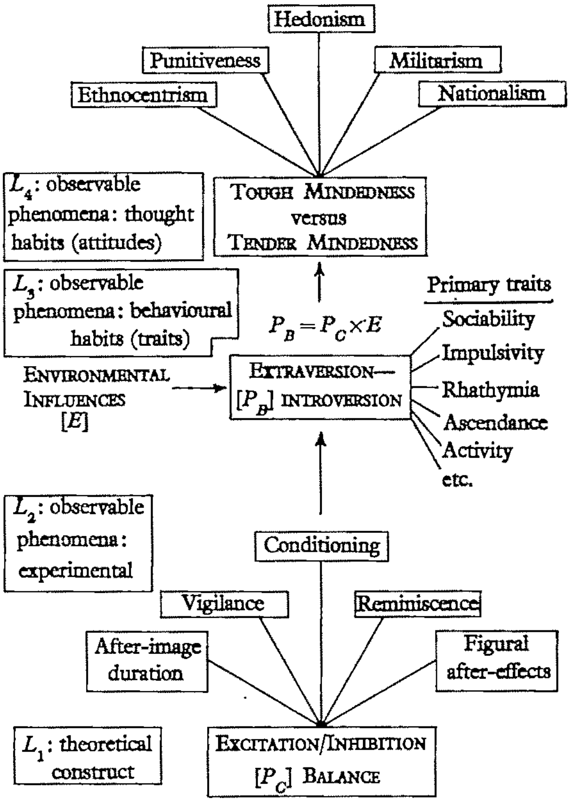

Hedonism
Punitiveness
Militarism
Ethnocentrism
Nationalism
L_4: observable phenomena: thought habits (attitudes)
TOUGH MINDEDNESS versus TENDER MINDEDNESS
L_3: observable phenomena: behavioural habits (traits)
$P_B = P_C \times E$
Primary traits
Sociability
Impulsivity
Rhathymia
Ascendance
Activity
etc.
ENVIRONMENTAL INFLUENCES [E]
EXTRAVERSION—[P_B] INTROVERSION
L_2: observable phenomena: experimental
Conditioning
Vigilance
Reminiscence
After-image duration
Figural after-effects
L_1: theoretical construct
EXCITATION/INHIBITION [P_C] BALANCE

Fig. 2. Diagram showing genotypic and phenotypic aspects of personality.

---

* It seems reasonable to suppose that genotypic differences will ultimately be linked up with observable structural differences by physiologists and neurologists; an attempt to frame certain hypotheses of a testable character along these lines has been made by Eysenck (1963*a*), who suggests that different parts of the ascending reticular formation may be implicated in the precise balance of the excitation/inhibition system. The effects of stimulant and depressant drugs on personality (Eysenck, 1963*b*) can also be brought into line by the assumption that the ascending reticular formation is concerned most intimately with the psychological constructs of excitation and inhibition.

We can now make deductions from these various postulates which enable us to perform critical experiments taking us out of the narrow circle of factor analysis altogether, and make possible the use of the much more powerful techniques of multiple discriminant function analysis. Consider the following experiment in which sixteen normal subjects, sixteen dysthymics and sixteen hysterics were given a battery of six tests, selected on the basis of the causal theory outlined above (Eysenck & Claridge, 1962). We can predict, of course, how each group shall score as compared with the others, but we can go further than that. Our theory predicts that, if we carry out a discriminant function analysis, this should give us two significant latent roots; it can further be predicted that if we derive variate scores for the forty-eight subjects of our experiment, they should be situated in a prescribed manner in a two-dimensional plane generated by the two significant variates. To put this prediction in its simplest form we may say that the mean variate scores for the three groups should lie at the corners of an equilateral triangle.

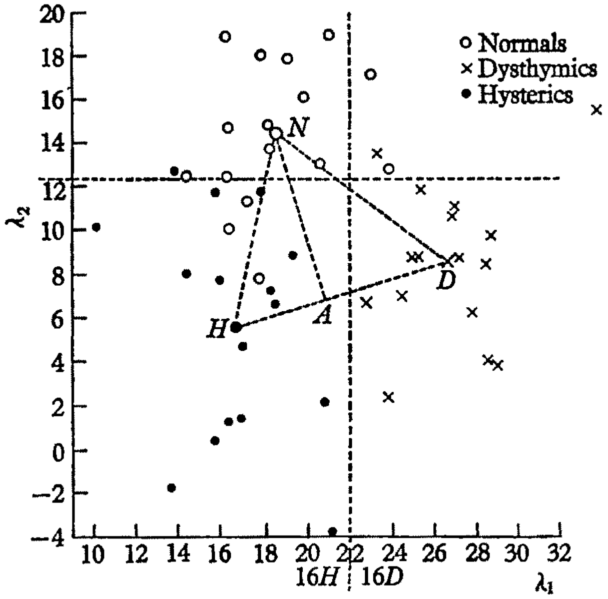


Fig. 3. Position of 16 normal, 16 dysthymic and 16 hysteric subjects in two-dimensional space generated by multiple discriminant function analysis.

Fig. 3 shows the outcome of the experiment. It will be seen that the prediction is verified, and that the first variate discriminates completely between the dysthymics and the hysterics. The second variate, with only slight overlap, discriminates between the normal group on the one hand, and the two neurotic groups on the other.

Even where a causal hypothesis is not available it is often possible to use discriminant function analysis to decide between two hypotheses regarding the description of personality. Consider two hypotheses very frequently advanced regarding the neurotic and psychotic disorders (Eysenck, 1955). Psychoanalysts often advocate the one-dimensional hypothesis; most psychiatrists, however, nowadays favour a two-dimensional hypothesis. A crucial test can, therefore, be devised involving the dimensionality of the performance of the three groups on a battery of tests selected on the basis of some hypothesis regarding their relevance to neurotic and psychotic disorder (Eysenck, 1955). In the actual experiment 20 normal controls, 20 neurotics and 20 psychotics were tested on four objective laboratory tests. Multiple discriminant function analysis disclosed two significant latent roots, thus rendering impossible the assumption that one dimension was sufficient to incorporate the results. Fig. 4 shows the actual positions of the members of the three groups; the correlation ratio between the three groups and the two variates was 0·84, which indicates a refreshingly high validity for the tests used in predicting these psychiatric criteria. That this figure is not higher is probably due to lack of reliability of the criteria; it will be seen in Fig. 4 that two of the neurotics, labelled *A* and *B*, were grouped with the psychotics by the tests. Both were readmitted later and diagnosed as psychotic.*

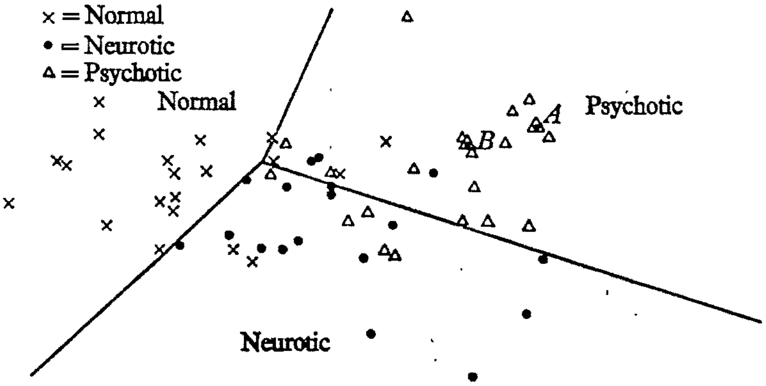


Fig. 4. Diagram showing position of 20 normal, 20 neurotic and 20 psychotic subjects in two-dimensional space generated by multiple discriminant function analysis.

There are other ways in which theories of this type can be tested. One of these is the *genetic* method. If it is true that psychotic and neurotic disorders are orthogonal to each other, then we would expect that the children of psychotic parents should not show any greater degree of neuroticism than would the children of normal parents. This very interesting hypothesis was tested by Cowie (1961), and her results leave no doubt that the genetic implication of neuroticism in the children of psychotic parents is non-existent; if anything they tended to be less neurotic! This finding may also serve as a warning to those who would overstress the importance of environment in giving rise to neurotic disorders; it is difficult to imagine a more severe stress to a child than having psychotic parents. In line with a generally hereditary view of the main dimensions of personality are also the results of a recent study of identical twins brought up in separation; in this work Shields (1962) found high correlations between the two twins for both extraversion and neuroticism; he also found that these correlations were, if anything, higher than corresponding ones for identical twins brought up together! This type of proof, which agrees well with previous studies by Eysenck & Prell (1951), Wilde (1962), Lienert & Reisse (1961), and many others, is relevant for the following reason. If we locate our axes in a random fashion, or according to some erroneous hypothesis, then we would not expect measures based on these placements to achieve any kind of biological reality. However, it has been amply demonstrated that extraversion, neuroticism, and psychoticism show a powerful, independent hereditary determination; it would seem to follow that the location of our axes cannot be random but must be at least to some degree in the right direction.

---

* It is interesting that cultural differences do not seem to affect the applicability of method or conclusion to any considerable extent. Devadasan (1963) has duplicated many of the details of S. B. G. Eysenck's (1956) study in this field on an Indian population in Kerala (Trivandrum) with almost identical results.

The last type of argument and proof, which I would suggest as appropriate, relates to the working out of aetiological models, and the design of methods of treatment related to these. It is a basic principle of behaviour therapy that neurotic disorders are simply maladaptive habits, acquired through a process of conditioning; or alternatively socially desirable habits which have failed to be acquired (Eysenck, 1960*c*, 1963*c*). This hypothesis has led to much work relating dysthymic disorders to overquick conditionability of patients, and hysteric and psychopathic disorders to chronic underconditionability of patients. (It will be remembered that overconditionability and underconditionability, respectively, are related to introversion and extraversion.) Support has already been brought forward to sustain these hypotheses, but I would be the last to claim that the case has, in any definitive sense, been proven; many points remain to be clarified and settled. The theory has, nevertheless, succeeded in giving rise to a method of treatment—behaviour therapy—which has been outstandingly successful as compared with previous methods. Again I will not claim too much for these new methods, and I will not go into the large and growing literature in any great detail, except to point out that success of treatment, if this is based on a definite theory, must to some degree strengthen the claim of that theory to be taken seriously. I would suggest, therefore, that aetiology and treatment must be taken into account in arriving at a final view of the adequacy of any principles of psychiatric classification claiming to be taken seriously.

The main points to emerge from this discussion are perhaps these. Factor analysis, principal component analysis, or some such technique, is necessary but not sufficient for the elaboration of a proper system of personality classification. The results achieved are inevitably subject to a large degree of subjectivity, and it is *in principle* impossible to avoid this subjectivity by statistical or mathematical manipulations. *The descriptive results of factor analysis require to be integrated with causal theories relating to the factors tentatively established or indicated.* It is only when these causal theories are tested and verified that the descriptive scheme can be accepted as forming part of the large body of data which make up experimental psychology.

There are, of course, many types of causal hypotheses which can be put forward in different situations, and there are many different types of deductions which can be made. It has been our task in this paper to touch in passing on several such causal hypotheses and deductions, and to show that the resulting picture is a reasonably congruent one which integrates observations, data, theories and experiments from a great variety of sources. There is, of course, no single proof of a theory such as the one here advocated, and no possibility of a crucial experiment; the burden of proof must lie in the general strength of the nomological network, linking together all these factors. It is believed that in this way psychiatric classification can be made much more reliable, valid and useful than it has been in the past when it relied exclusively on subjective observation, non-quantitative argument and non-experimental demonstration. Obviously the procedure of making our typologies more scientific has only just begun, and still has a long way to go before we can hope to achieve a satisfactory level of accuracy, reliability, and validity; nevertheless, the success which has attended our first faltering steps does suggest that the method followed is the correct one and will in due course lead to a better understanding as well as to a better description of human behaviour and personality.

REFERENCES

ANDERSON, H. H. (1939). Domination and integration in the behaviour of kindergarten children and teachers. *Genet. Psychol. Monogr.* **21**, no. 3.

BARTLETT, F. C. (1932). *Remembering.* Cambridge University Press.

BARTLETT, M. S. (1948). Internal and external factor analysis. *Brit. J. Stat. Psychol.* **1**, 73–81.

BARTLETT, M. S. (1950). Tests of significance in factor analysis. *Brit. J. Stat. Psychol.* **3**, 77–85.

BERNARD, C. (1878). *Leçons sur les phénomènes de la vie communs aux animaux et aux végétaux.* Paris: Baillière.

BINDRA, D. & SPINNER, N. (1958). Response to different degrees of novelty; the incidence of various activities. *J. Exp. Anal. Behaviour,* **1**, 341–50.

BROWN, F. W. (1942). Heredity in the psychoneuroses. *Proc. R. Soc. Med.* **35**, 785–90.

CATTELL, R. B. (1952). *Factor Analysis.* New York: Harper.

CATTELL, R. B. & SCHEIER, I. H. (1961). *The Meaning and Measurement of Neuroticism and Anxiety.* New York: Ronald.

CLARIDGE, G. S. (1960). The excitation-inhibition balance in neurotics. In *Experiments in Personality,* vol. II. Edited by H. J. Eysenck. London: Routledge and Kegan Paul.

CLARIDGE, G. S. (1961). Arousal and inhibition as determinants of the performance of neurotics. *Brit. J. Psychol.* **52**, 53–63.

CLARIDGE, G. S. & HERRINGTON, R. N. (1960). Sedation threshold, personality, and the theory of neurosis. *J. Ment. Sci.* **106**, 1568–83.

COHEN, M. E. (1951). The high familial prevalence of neurocirculatory asthenia (anxiety neurosis, effort syndrome). *Amer. J. Hum. Genet.* **3**, 126–58.

COWIE, V. (1961). The incidence of neurosis in the children of psychotics. *Acta Psychiat. Scand.* **37**, 37–87.

DENCKER, S. J. (1958). *A Follow-up Study of* 128 *Closed Head Injuries in Twins Using Co-twins as Controls.* Copenhagen: Munksgaard.

DEVADASAN, K. (1963). Personality dimensions: a critical study. Ph.D. Thesis. University of Kerala, India.

EHRENBERG, A. S. C. (1962). Some questions about factor analysis. *The Statistician,* **12**, 191–208.

EYSENCK, H. J. (1947). *Dimensions of Personality.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1950). Criterion analysis—an application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.* **57**, 38–53.

EYSENCK, H. J. (1952*a*). *The Scientific Study of Personality.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1952*b*). Schizothymia-cyclothymia as a dimension of personality. II. Experimental. *J. Personality,* **20**, 345–84.

EYSENCK, H. J. (1952*c*). Uses and abuses of factor analysis. *Appl. Statistics,* **1**, 45–9.

EYSENCK, H. J. (1953). The logical basis of factor analysis. *Amer. Psychologist,* **8**, 105–14.

EYSENCK, H. J. (1954*a*). A psychological approach to the problem of non-uniqueness in multivariate solutions. *Proc. Conference on Multi-Dimensional Analysis,* 24–7. Ann Arbor: University of Michigan Press.

EYSENCK, H. J. (1954*b*). Zur Theorie der Persönlichkeitsmessung. *Z. diag. Psychol. Persönlichkeitsf.* **2**, 87–101, 171–87.

EYSENCK, H. J. (1954*c*). Abord statistique et experimental du problème typologique dans la personalité neurotique, psychotique et normale. *L' Evolution psychiatrique,* **3**, 377–404.

Eysenck, H. J. (1955). Psychiatric diagnosis as a psychological and statistical problem. *Psychol. Rep.* **1**, 3–17.

Eysenck, H. J. (1956*a*). L'analyse factorielle et le problème de la validité. *Colloques int. Cent. nat. Rech. sci., Paris*, **58**, 237–52.

Eysenck, H. J. (1956*b*). The questionnaire measurement of neuroticism and extraversion. *Riv. Psicol.* **50**, 113–140.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

Eysenck, H. J. (1959). *Manual of the Maudsley Personality Inventory.* London: University of London Press.

Eysenck, H. J. (1960*a*). *The Structure of Human Personality.* London: Methuen.

Eysenck, H. J. (1960*b*). Levels of personality, constitutional factors, and social influences: an experimental approach. *Int. J. Soc. Psychiat.* **6**, 12–24.

Eysenck, H. J. (1960*c*). *Behaviour Therapy and the Neuroses.* Oxford: Pergamon.

Eysenck, H. J. (1960*d*). *Experiments in Personality.* Edited by H. J. Eysenck. London: Routledge and Kegan Paul.

Eysenck, H. J. (1960*e*). Classification and the problem of diagnosis. In *Handbook of Abnormal Psychology*. Edited by H. J. Eysenck. London: Pitman Medical.

Eysenck, H. J. (1961). Personality and social attitudes. *J. Soc. Psychol.* **53**, 243–8.

Eysenck, H. J. (1962). Correspondence. *Brit. J. Psychol.* **53**, 455–6.

Eysenck, H. J. (1963*a*). Biological basis of personality. *Nature, Lond.*, **199**, 1031–4.

Eysenck, H. J. (1963*b*). *Experiments with Drugs.* Edited by H. J. Eysenck. Oxford: Pergamon.

Eysenck, H. J. (1963*c*). *Experiments in Behaviour Therapy.* Edited by H. J. Eysenck. Oxford: Pergamon.

Eysenck, H. J. & Claridge, G. (1962). The position of hysterics and dysthymics in a two-dimensional framework of personality description. *J. Abnorm. Soc. Psychol.* **69**, 46–55.

Eysenck & Prell, P. (1951). The inheritance of neuroticism; an experimental study. *J. Ment. Sci.* **97**, 441–65.

Eysenck, S. B. G. (1956). Neurosis and psychosis: an experimental analysis. *J. Ment. Sci.* **102**, 517–29.

Eysenck, S. B. G., Eysenck, H. J. & Claridge, G. (1960). Dimensional personality, psychiatric syndromes, and mathematical models. *J. Ment. Sci.* **106**, 581–9.

Foulds, G. A. (1961). The logical impossibility of using hysterics and dysthymics as criterion groups in the study of introversion and extraversion. *Brit. J. Psychol.* **52**, 385–7.

Foulds, G. A. (1962*a*). A quantification of diagnostic differentia. *J. Ment. Sci.* **108**, 389–405.

Foulds, G. A. (1962*b*). Correspondence. *Brit. J. Psychol.* **53**, 456–7.

Franks, C. M., Souieff, M. I. & Maxwell, A. E. (1960). A factorial study of certain scales from the MMPI and the STDCR. *Acta Psychol.* **17**, 407–16.

Gray, J. (1964). *Pavlov's Typology: Recent Theoretic and Experimental Works from the Laboratory of B. M. Teplov.* Edited by J. Gray. Oxford: Pergamon.

Hallgren, B. (1950). *Specific Dyslexia. A Clinical and Genetic Study.* Stockholm: Esselte aktiebolag.

Hallgren, B. (1957). *Enuresis. A Clinical and Genetic Study.* Stockholm: Esselte aktiebolag.

Hildebrand, H. P. (1958). A factorial study of introversion-extraversion. *Brit. J. Psychol.* **49**, 1–11.

Hope, K. (1963). The structure of hostility in a sample of normal and neurotic persons. Ph.D. thesis, University of London.

Ingham, J. G. (1962). Correspondence. *Brit. J. Psychol.* **53**, 458–9.

Jeffers, J. N. R. (1962). Principal component analysis of designed experiment. *The Statistician*, **12**, 230–42.

Kaij, L. (1960). *Alcoholism in Twins. Studies on the Etiology and Sequels of Abuse of Alcohol.* Stockholm: Almqvist and Wiksell.

Kendall, M. G. (1950). Factor analysis as a statistical technique. *J. Roy. Stat. Soc.*, Vol. 12, Series B, 60–73.

Kendall, M. G. (1957). *A Course in Multivariate Analysis.* London: Griffin.

Kestelman, H. (1952). The fundamental equation of factor analysis. *Brit. J. Stat. Psychol.* **5**, 1–6.

KNAPP, R. R. (1962). *The Maudsley Personality Inventory: Manual.* San Diego, California: Educational and Industrial Testing Service.

KRAULIS, W. (1931). Zur Vererbung der hysterischen Reaktionsweise. *Z. ges. Neurol. Psychiat.* **136**, 174–258.

KSHIRSAGAR, A. M. (1961). The goodness of fit of a single non-isotropic hypothetical principal component. *Biometrika*, **48**, 397–407.

LAWLEY, D. N. & MAXWELL, A. E. (1962). Factor analysis as a statistical method. *The Statistician*, **12**, 209–29.

LEWIS, A. (1935). Problems of obsessional illness. *Proc. R. Soc. Med.* **29**, 325–36.

LEWIS, A. J. & MAPOTHER, E. (1941). In *Price's Textbook of the Practice of Medicine.* London: Oxford University Press.

LIENERT, G. & REISSE, H. (1961). Ein korrelativer-analytischer Beitrag zur genetischen Determination des Neurotismus. *Psychol. Beiträge*, **7**, 121–30.

LJUNGBERG, L. (1957). Hysteria. A clinical, prognostic and genetic study. *Acta psychiat. neurol. Scand.*, Suppl. 112.

LUXENBURGER, H. (1930). Heredität und Familientypus der Zwangsneurotiker. *Arch. Psychiat.* **91**, 590–4.

MCCLELLAND, D. C. (1951). *Personality.* New York: Sloane.

MCGUIRE, R. J., MOWBRAY, R. M. & VALLANCE, R. C. (1963). The Maudsley Personality Inventory used with psychiatric inpatients. *Brit. J. Psychol.* **54**, 157–66.

MCINNES, R. G. (1937). Observations on heredity in neurosis. *Proc. R. Soc. Med.* **30**, 895–904.

MARTIN, D. V. & CAINE, T. M. (1963). Personality change in the treatment of chronic neurosis in a therapeutic community. *Brit. J. Psychiat.* **109**, 267–72.

RÜDIN, E. (1953). Ein Beitrag zur Frage der Zwangskrankheit, insbesondere ihrer hereditären Beziehungen. *Arch. Psychiat.* **191**, 14–54.

SHAGASS, C. & NAIMAN, J. (1956). The sedation threshold as an objective index of manifest anxiety in psychoneurosis. *J. Psychosom. Res.* **1**, 49–57.

SHIELDS, J. (1954). Personality differences and neurotic traits in normal twin shool children. *Eugen. Rev.* **45**, 213–45.

SHIELDS, J. (1962). Monozygotic twins brought up apart and brought up together. *An Investigation into the Genetic and Environmental Causes of Variation in Personality.* London: Oxford University Press.

SIGAL, J. J., STAR, K. H. & FRANKS, C. M. (1958). Hysterics and dysthymics as criterion groups in the study of introversion-extraversion. *J. Abnorm. Soc. Psychol.* **57**, 143–8.

SLATER, E. (1943). The neurotic constitution. *J. Neurol. Psychiat.* **163**, 1–16.

SLATER, E. (1961). Hysteria 311. *J. Ment. Sci.* **107**, 359–81.

SLATER, E. & SLATER, P. (1944). A heuristic theory of neurosis. *J. Neurol. Neurosurg. Psychiat.* **7**, 49–55.

SLATER, P. (1944). *The Selective Vocabulary Test.* London: Harrap.

SLATER, P. (1953). The factor analysis of matrices of negative correlations. *Brit. J. Stat. Psychol.* **6**, 101–6.

SLATER, P. (1960*a*). The reliability of some methods of multiple comparison in psychological experiments. Ph.D. thesis, University of London.

SLATER, P. (1960*b*). The analysis of personal preferences. *Brit. J. Stat. Psychol.* **13**, 119–35.

SPEARMAN, C. (1923). *The Nature of Intelligence and the Principles of Cognition.* London: Macmillan.

THOMSON, G. H. (1951). *The Factorial Analysis of Human Ability*, 5th ed. London: University Press.

THURSTONE, L. L. (1947). *Multiple Factor Analysis.* Chicago: University Press.

VERNON, P. E. (1953). *Personality Tests and Assessments.* London: Methuen.

WARBURTON, F. W. (1954). The full factor analysis. *Brit. J. Stat. Psychol.* **7**, 101–6.

WILDE, G. J. S. (1962). *Neurotische Labiliteit gemeten volgens de Vragenlijstmethode.* Amsterdam: Uitg. F. van Rossen.

WUNDT, W. (1903). *Grundzüge der physiologischen Psychologie*, 5th ed. vol. 3. Leipzig: W. Engelmann.

(*Manuscripts received May/November* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 295–301
Printed in Great Britain

RETROACTIVE INTERFERENCE IN THE RECALL OF SIMPLE SENTENCES

BY JACQUES MEHLER AND GEORGE A. MILLER
Harvard University

A list of English sentences was learned by the method of free recall. Performance was scored both for sentences recalled precisely and for sentences recalled in content regardless of their syntactic form. Interpolated lists of sentences differing from the original list in syntactic details or in semantic form were used to produce retroactive inhibition (RI). The RI for semantic content was negligible; when the interpolated list differed only in syntax and preserved the semantic content of the original list the effect was facilitatory rather than inhibitory. When performance was scored for syntactic as well as semantic aspects RI was appreciable. The results are interpreted as additional support for the hypothesis that the two aspects, semantic content and syntactic form, pose more or less distinct tasks for the learner.

In previous articles (Miller, 1962; Mehler, 1963) it has been argued, on the basis of errors observed in the free recall of sentences varying in both content and grammatical form, that when a person tries to memorize a sentence he deals with it in (at least) two separate parts—one carrying its meaning, or semantic content, the other carrying information about its syntactic structure. The present paper reports a further attempt to test the validity of that general distinction.

Since retroactive inhibition (RI) is ordinarily easier to demonstrate with nonsense syllables than with meaningful materials, and since, under the experimental conditions employed in these studies, the syntactic structure is relatively arbitrary, whereas the content of the sentence is meaningful, it seemed plausible that RI might affect these two aspects of the learner's task differentially.

The terminology of transformational grammar (Chomsky, 1957) will be used to describe the syntactic information that has to be remembered. The assumption is that most sentences are derived from a more fundamental type of sentence by certain syntactic rules called 'optional transformations'. The fundamental sentences, or *k-sentences* (derived from underlying kernel strings using only obligatory transformations), are simple, active, affirmative, declarative sentences, e.g., *The boy has hit the ball.* Three singular optional transformations are used in this experiment: the negative (N), the passive (P), and the interrogative (Q). Application of these three transformations in all possible combinations can produce seven additional sentences from each k-sentence (K).

The relations among these eight sentences can be represented by a cube, as in Fig. 1. Each vertex of the cube represents a separate syntactic form; each edge corresponds to a pair of sentences that differ only by a single transformation.

EXPERIMENTAL PROCEDURE

A familiar RI paradigm was used; original learning (OL) was followed by controlled intervening activity (IL), then by a retest (T) on the original material. The interpolated tasks differed for different groups of subjects, but the original learning and final test were the same for all.

Sets of eight sentences were presented to be learned by the method of free recall. Each sentence was typed on a separate 3 in. × 5 in. card and one card was presented manually every 4 sec.

Subjects were instructed to remember each sentence as precisely as they could, because a recall test would be given when the deck of cards was finished. After all eight sentences had been presented, the subject was asked to write down, in any order, as many sentences as he could remember. This procedure was repeated for a total of either two or four learning trials, depending on the experimental condition to which the subject had been assigned; each learning trial was followed immediately by a test of recall. The order of presentation of the sentences was varied for each trial.

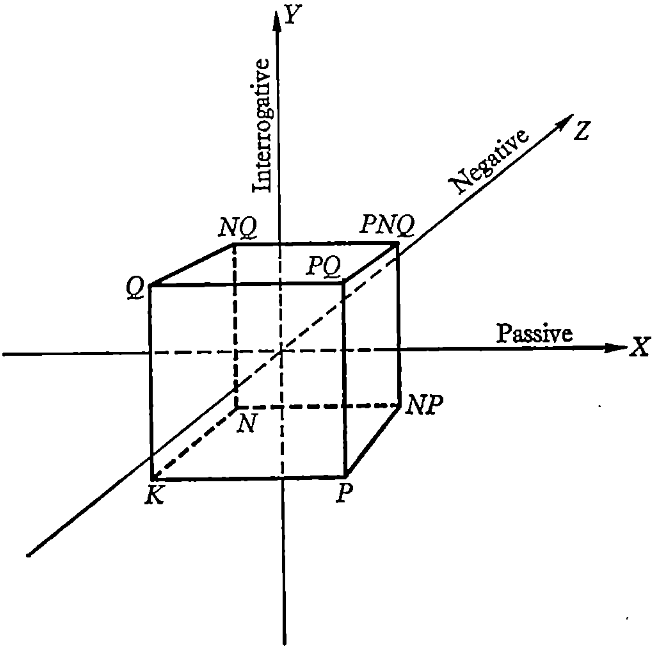


Fig. 1. Representation of the eight syntactic forms in terms of the corners of a cube, when the co-ordinates are labelled with the names of grammatical transformations.

There were three different interpolated tasks, one designed to produce syntactic interference, a second designed to produce semantic interference, and a third (control) intended to interfere minimally with the eight sentences originally learned. Each subject had only one kind of interpolated task. All of the experimental groups had four trials of IL.

Immediately after the interpolated task was completed, each subject was asked to recall all he could remember of the original eight sentences. When the subject reported that he had recalled as much as he could, he was told that if he thought that with more time he could do better, he could take two or three minutes more.

The 100 subjects used in this study were all undergraduates at Radcliffe College or Harvard University; they were paid $1.50 per hour for their services.

Materials

The following list of eight sentences was used as the original learning task for all groups:

The boy has taken the photograph.
The man hasn't bought the house.
The passenger has been carried by the airplane.
Has the girl worn the jewel?
The discovery hasn't been made by the biologist.
Hasn't the secretary typed the letter?
Has the tree been hit by the car?
Hasn't the essay been written by the student?

The syntactic form of these sentences will be referred to respectively as K, N, P, Q, PN, NQ, PQ and NPQ.

Syntactic interference. Ten of the subjects in this group had two trials to learn the original material; thirty subjects had four trials. The interpolated material was derived by modifying the syntactic structure of the sentences learned originally.

The thirty subjects having four trials of OL were divided into three subgroups of ten each; the subgroups differed only with respect to the degree of syntactic similarity between the original and interpolated materials. The interfering materials were derived from the original list by rotating the cube in Fig. 1. For the first subgroup, each sentence in the intervening list differed by only one transformation from its corresponding sentence on the original list. Of the three lists that it is possible to derive in this way, only the one obtained by rotation around the Y-axis was used. For the second subgroup, each sentence in the intervening list differed by exactly two transformations from its corresponding sentence in the original list; the derived list was obtained by rotation of the original two steps around the Y-axis. And in the third subgroup, each sentence differed in all three transformations. The syntactic changes introduced in order to produce the interpolated lists are given explicitly in Table 1. These transformed sentences were presented for four trials in the same way as the original sentences had been.

Table 1. *Syntactic changes used to produce three types of interpolated lists differing only in syntactic form from the original list*

	Interpolated list		
Original list	1 step	2 step	3 step
K	P	PN	PNQ
N	K	P	PQ
P	PN	N	NQ
Q	PQ	PNQ	PN
PN	N	K	Q
NQ	Q	PQ	P
PQ	PNQ	NQ	N
PNQ	NQ	Q	K

Semantic interference. Ten of the subjects in this group had two trials to learn the original material; twenty subjects had four trials. All thirty subjects were then asked to learn a new set of eight sentences, completely unrelated to their original learning. These sentences, presented for four trials, in exactly the same way as the sentences that were learned originally, were:

The doctor has treated the patient.
The child hasn't eaten the pie.
The money has been stolen by the burglar.
Has the candidate won the election?
The pipe hasn't been corroded by the water.
Hasn't the horse jumped the fence?
Has the deer been shot by the hunter?
Hasn't the card been punched by the machine?

It will be noted that these eight sentences, like the original ones, included one representative of each of the eight syntactic structures represented in the cube in Fig. 1, but their semantic content is, of course, different.

Control. Ten of the subjects in the control group had two trials to learn the original material and twenty had four trials. Then, instead of learning a new list, these subjects added columns of 4-digit numbers (taken from a table of random numbers) for a period of 13 min. each. Subjects were told that they were being tested for mathematical ability, and were asked to make as few errors as possible. They were also required to keep a running total, since they were told that they might be stopped at any point and that their score would be the number of digits they had correctly added to that point.

SCORING PROCEDURES

Performance in this situation can be scored in several ways; two of them have been used and are described here.

Consider first the *overall sentence score*. A response sentence can, of course, be scored as completely correct only if it is a verbatim reproduction of a stimulus sentence. Any difference between the stimulus sentence and the recalled sentence can be counted as an error; the only exception (which occurred infrequently) was that the substitution of an obvious synonym was not scored as an error. These overall sentence scores constitute the basic data from the experiment.

A second type of score was used in order to distinguish between semantic and syntactic sources of error. We shall say that two sentences are derived from the same kernel strings if we can obtain one from the other by applying one or more optional transformations (or their inverses) to one member of the pair. (As before, if one word in one sentence is a synonym of the equivalent in the other sentence, the pair is still considered to have the same sentence content if the above definition holds.) Two sentences that are derived from the same kernel are said to share the same semantic content and to differ only in their syntactic features.

This definition enables us to separate the errors into two classes, semantic and syntactic, and so to define a *sentence-content* score. In this method of scoring, a response is considered correct if it shares the semantic content of the stimulus sentence, i.e., if it is correct overall or if it can be obtained from the stimulus sentence by the singularly optional transformations N, P, and Q or their inverses (or by synonym substitution where the case might arise).

Percentage RI was computed by the formula $100\ (T_c - T_e)/T_c$, where T_c and T_e are the retest scores for the control and experimental groups, respectively.

RESULTS

Before considering the results for the main experiment, let us first dispose of the three subgroups subjected to alternative types of syntactic interference. In the present experiment one might assume that the three rotations of Table 1 define equal intervals along some scale of syntactic similarity; since it is well known that the similarity of the original and interpolated materials is an important variable affecting RI, we might expect to find that RI is a function of syntactic similarity, so defined.

From the top half of Table 2, however, it can be seen that although the amount of overall RI is fairly large for all three subgroups, the differences among them are relatively small and not statistically significant ($\chi^2 = 2{\cdot}7$). The differences that do exist may be attributable to differences in the levels of mastery of OL. In any case, there seemed to be no significant trend, so the results for the three subgroups were pooled into a single group of thirty subjects, hereafter called the syntactic interference group.

Now return to the main experiment, which was designed to test the hypothesis that, in the context of this experiment, sentence content is more resistant to RI than is syntactic form. If we consider first the overall scores in the lower half of Table 2, we see that syntactic interference produced more RI (48%) than did semantic interference (18 %), presumably because of the greater similarity between OL and IL for the syntactic group. In part, however, this difference is attributable to the difference in the level of mastery on the final (fourth) trial of OL; when we apply a constant of proportionality to equate the syntactic group to the other two groups on OL, the estimate of RI is reduced somewhat for the syntactic group, but is still larger than for the group exposed to semantic interference.

Table 2. *Average overall and content scores (maximum score = eight sentences) on the final (fourth) trial of OL (4 OL) and on the retest (T) after IL, and percentage RI, for groups having syntactic, semantic, and neutral IL*

		Overall score			Content score		
Condition	(*N*)	4 OL	T	% RI	4 OL	T	% RI
Syntactic interference							
1 step	10	4·4	2·5	55	7·8	7·8	2
2 step	10	4·8	3·9	30	7·3	7·4	7
3 step	10	4·4	2·4	57	7·7	7·5	6
Syntactic interference combined	30	4·5	2·9	48	7·6	7·6	5
(Corrected*)		(5·9)	(3·8)	(32)			
Semantic interference	20	5·8	4·6	18	7·5	7·0	13
Control group	20	6·0	5·6	—	7·5	7·9	—

* Corrected scores were obtained by multiplying observed scores by a constant of proportionality to equate this group with the others on OL.

When we consider the content scores, however, we see that there is less RI than for the overall scores, which would seem to provide evidence in support of the two-aspect hypothesis that the experiment was designed to test. This conclusion is questionable, however, on the grounds that by the fourth trial all the groups had learned the eight sentence contents almost perfectly, which may somehow have insulated them against the effects of IL. Certainly, if the hypothesis was correct, we should expect to find some facilitation in the content scores for the group with syntactically interfering IL, since the IL simply repeated the same sentence contents that had been presented during OL. Performance was so near perfect, however, that no such facilitation could be observed.

In order to get a clearer picture of the effects of IL on the content scores, therefore, we must consider the data reported in Table 3. Here the conditions are exactly the same, except that only two trials of OL (rather than four) were given. From the content scores in Table 3 we see that there was indeed facilitation, as expected, for the syntactic group. Indeed, the facilitation was sufficient to produce an apparent facilitatory effect in the overall scores for this group, in spite of the similarity of OL and IL. Again, however, there is an unfortunate sampling error that produced a difference in the level of mastery of OL; when the correction is made, as before, in the results for the group with syntactic interference, the magnitude of this facilitation is markedly reduced.

Table 3 also indicates that for the group with semantic interference RI was greater (38 %) when both syntactic and semantic features were scored than when only the semantic aspects were considered (14 %).

No intrusion of sentence contents from either list was observed with the semantic interference group.

Table 3. *Average overall and content scores (maximum score = eight sentences) on the final (second) trial of OL (2 OL) and on the retest (T) after IL, and percentage RI, for groups having syntactic, semantic, and neutral IL*

		Overall score			Content score		
Condition	*N*	2 OL	T	% RI	2 OL	T	% RI
Syntactic interference	10	3·7	2·8	−33	6·3	7·3	−49
(Corrected*)		(2·6)	(2·0)	(+5)	(5·2)	(6·0)	(−22)
Semantic interference	10	2·6	1·3	38	5·0	4·2	14
Control group	10	2·6	2·1	—	5·4	4·9	—

* Corrected scores were obtained by multiplying observed scores by a constant of proportionality to equate this group with the others on OL.

Discussion

The results indicate that there is greater RI, under the conditions of these experiments, for the syntactic than for the semantic aspects of a sentence. This conclusion would appear to be consistent with the hypothesis that subjects deal with sentences on two distinct levels; the semantic aspects are coded separately and are relatively free of RI, whereas the syntactic aspects are (in this situation) much more arbitrary, harder to remember, and more subject to RI. The alternative hypothesis, that each sentence is an arbitrary string of symbols to be memorized without regard for syntactic form or semantic content, would be not adequate to predict the differences here obtained between the two methods of scoring.

The principal puzzle posed by these results appears in the overall scores for the group subjected to syntactic interference during IL. Why is the syntactic interference ineffective (or even facilitatory) after two trials of OL, but inhibitory after four trials? If formal similarity between OL and IL were the only factor at work, one would expect strong RI in both experiments.

This puzzle provokes the following speculation. If we accept the two-aspect hypothesis, it is but a short step to a two-stage hypothesis about the process of memorizing sentences. That is to say, all the evidence to date is consistent with the notion that subjects first learn the sentence content, then learn the syntactic details. According to this line of argument, during the first two trials our subjects were concentrating on remembering what the eight sentences were about, so an interpolated task that repeated those sentence contents would scarcely interfere, and might even prove helpful. By the fourth trial, however, subjects had mastered the contents of the sentences and were well into the second stage, trying to straighten out the syntactic details, so an interpolated task that repeated those contents would be of little help, and the scrambled syntax could cause considerable confusion. This argument would explain the major discrepancy between the data of Tables 2 and 3.

This investigation was supported by funds granted to Harvard University, Centre for Cognitive Studies, from the National Institutes of Health, United States Public

Health Service (PHS Research Grant No. MH–05120–02); from the National Science Foundation (Grant No. NSF G-16486); and from the Carnegie Corporation of New York.

REFERENCES

CHOMSKY, N. (1957). *Syntactic Structures.* The Hague: Mouton.

MEHLER, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *J. Verb. Learn. Verb. Behav.* **2,** 346–51.

MILLER, G. A. (1962). Some psychological studies of grammar. *Amer. Psychol.* **17,** 748–62.

(*Manuscript received* 21 *August* 1963)

Brit. J. Psychol. (1964), 55, 3, pp. 303–306
Printed in Great Britain

RATE OF PRESENTATION AND ORDER OF RECALL IN IMMEDIATE MEMORY

By MICHAEL I. POSNER
University of Wisconsin

University students attempted immediate recall of 8-digit series presented aurally. When subjects were required to report the digits in their presented order, the fast presentation rate (96 digits per minute) was superior to the slow rate (30 per minute). When subjects were required to report the last four digits, before giving the first four digits, this improved performance at the slow rate but not at the fast rate. These data resolve previous conflicting results and tend to support a decay factor in immediate memory.

I. Introduction

Varying the rate of presentation of items for immediate recall serves two functions. First, it changes the time that an item must stay in store prior to recall and, secondly, it alters the time for rehearsal, association and other strategies to be brought into play. Thus Posner (1963) has suggested that increasing the rate of presentation will only show improved recall because of decreased time in store in tasks which tend to reduce the use of recall strategies.

A number of studies have manipulated time in store by varying the rate of presentation of aural sequences of digits for immediate recall. Conrad & Hille (1958) and Fraser (1958) have shown improved recall with increased rate of presentation. Many studies which appear to be similar have shown no differences or even the opposite effect of rate (Calhoon, 1935; Fraisse, 1945; Peatman & Locke, 1934; Pollack, 1952). The Conrad & Hille (1958) study used several conditions in which the recall itself was rigidly paced but also contained a condition using unpaced recall which was superficially identical with those of Pollack (1952), yet the results were opposite. One reason for these discrepancies may involve the order in which subjects actually record the digits. As Mackworth (1962*a*) has pointed out, in a situation using written recall where only the final order need be correct some subjects will write down the last digits first. Conrad & Hille (1958), however, only tested unpaced recall after several conditions in which pacing forced the order of recall to match the order of presentation. These previous conditions may have set subjects to record the digits in the order received. On the other hand, studies finding opposite results to those of Conrad & Hille (1958) were run in groups and report no effort to control the actual order in which messages were recorded.

Since transformations of order involve a recall strategy which may be affected by rehearsal it is possible that allowing subjects to record the digits in an order different from the presentation would be more beneficial at slow rates. The present study was designed, therefore, first to check whether order of recall interacts with rate of presentation in a manner which could cause these varying results and, secondly, to test the effect of rate upon recall with order fixed.

II. METHOD

Subjects. Twenty subjects were used in each of the main experiments and nine in the control group. All were students in introductory psychology courses of the University of Wisconsin and received course points for their participation.

Materials. Ten practice and forty regular series were recorded at each of two speeds, 30/min. and 96/min. Each series consisted of eight digits selected at random. The numbers in the series were recorded over white noise which served to reduce distractions from outside the experiment. There was a 20 sec. pause after each series.

Procedure. In Expt. I twenty subjects were run individually. They were randomly assigned to one of two groups. Group F received the series at the fast rate and group S at the slow rate. Subjects in each group performed two different tasks. In task O, after hearing all eight digits, the subjects repeated them back to the experimenter in the order presented. In task R they were required to present digits 5–8 first and then 1–4. To determine the order of the oral recall subjects said 'blank' to indicate omitted digits. Within each group, half the subjects performed twenty series of condition O first and then, following 30 min. of an unrelated task, recalled twenty series in condition R; the other half did the reverse.

Expt. II was the same as described above except that the subjects heard the numbers through earphones instead of over a loudspeaker. In this experiment subjects wrote their recall on a score sheet which contained eight boxes per series clearly divided into two groups of four. During the practice series subjects were instructed to record the answers only in the way indicated by the particular condition being studied. Subjects were checked to be sure they followed this procedure.

Nine subjects were run in a group as a control. They listened to the digits played over the speaker, as in Expt. I, but they wrote the digits as presented rather than having to recall them. The subjects received 20 series at the slow speed followed by 20 at the fast speed.

Scoring. The basic score used was percentage of correct series. If any digit was incorrect the series was marked wrong. To compute serial position curves the number of correct digits at each position was computed. If any but the correct digit was inserted in a given position it was scored as an error.

III. RESULTS

The basic results of the two experiments are shown in Table 1. The mean difference between the two tasks (R – O) is significantly higher at the slow speed in both studies. The results seem to show conclusively that the inverted recall of task R greatly improved performance at the slow speed while tending to have the opposite effect at the fast rate.

Table 1. *Mean percentage of correct series and significance tests*

	Fast	Slow	t
Task O			
Expt. I	51·5	29·5	2·13*
Expt. II	66·5	47	1·6
Task R			
Expt. I	52	50	0·2
Expt. II	59·5	60·5	0·08
Mean difference	0·5	20·5	2·26*
(R – O)	–7	13·5	3·14**

* $P < 0·05$; ** $P < 0·01$.

Looking at the differences in performance at the two rates, it is clear that if a reversal (task R) is required there are no differences between the two speeds. This is a frequent result of experiments which do not control order of recall. If, however, only the order condition is allowed, the fast speed appears to show better performance. In Expt. I the mean difference in recall between speeds is 22% ($P < 0{\cdot}05$), while in the second experiment the difference is almost as great, 19·5%, but is not significant. Since this comparison is between independent groups at the two speeds the variability is quite high. However, the overall results seem to indicate that with the order condition required the fast rate is superior, as would be predicted by Conrad & Hille (1958).

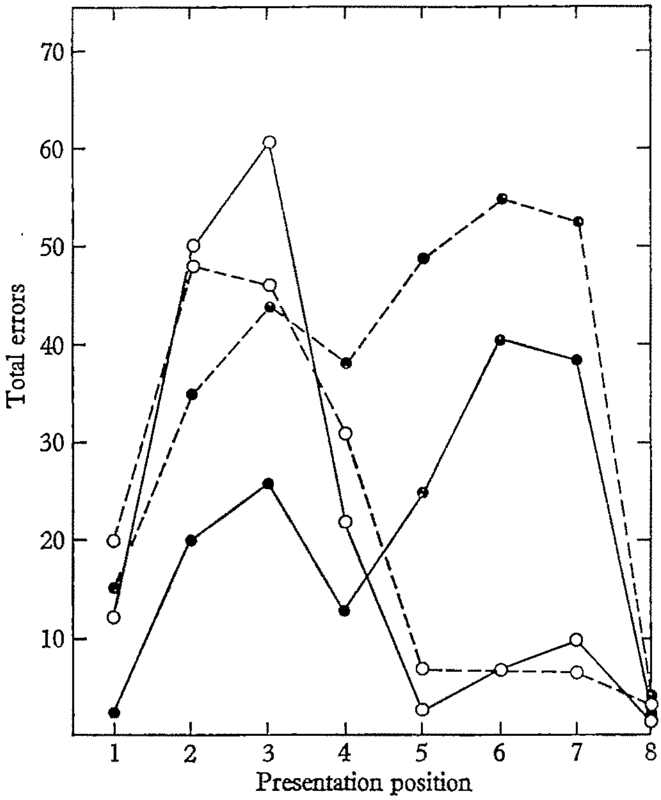


Fig. 1. Serial position curves, Expt. II. ●, Order; ○, reverse; —, fast; - - -, slow.

To get a more complete idea of the effect of the transformation required in the reverse condition, Fig. 1 shows the serial position curves for Expt. II. (The total possible error score for any one digit is 200.) Despite the use of oral recall, the serial position curves of Expt. I are virtually identical with those presented here. In the order condition recall at the fast rate is superior at every serial position, while in the reverse condition there is little difference between the two rates at any of the positions. It is clear that at both speeds prior recall of digits 5–8 improves performance at these positions, but at the fast rate this is accompanied by a much greater loss in retention at the first four positions than at the slow rate. This result would seem to agree with subjects' reports that, with presentation at the slow rate, rehearsal during the first four items preserves them during the subsequent presentation and recall of the last four. It may also be argued that reversal at the fast rate is difficult because of the problem of recognizing the fifth digit, but the serial position curves give no support to this explanation since the fifth digit at the fast rate is virtually errorless.

Two additional results should be briefly outlined. The increase in recall in all conditions of Expt. II over Expt. I seems to be explicable on the basis of the increased control over distraction which the earphones provided. It has also been shown previously that different types of responses may cause varying amounts of interference (Mackworth, 1962*b*). It is possible that written responses may cause less interference than oral responses. The control group was run to check the ability of the subjects to hear and record the digits when memory was not required. The results in terms of mean percentage of correct series were 98·3 for the slow rate and 95 for the fast rate. This indicates that whatever differences in ability to hear the series were present would not account for the superior performance at the fast rate found in the main studies.

IV. DISCUSSION

This experiment illustrates at least one reason for differences in results among studies of the effect of presentation rate upon immediate recall. If experiments do not control the order in which subjects actually record the responses they tend to show no great differences in performance between slow and fast rates. It is not surprising that the slow rate is actually superior in some studies, since recalling the last four digits first is not necessarily an optimal strategy and in free conditions even more successful transformations might be used. If, however, order of recall is controlled, this study agrees with those finding better recall at fast rates.

Mackworth (1962*a*) has also shown different serial order curves in her visual task with different orders of recall, although she does not report whether the subjects required to recall in order differed from her overall findings of improved recall with decreasing rate. Therefore it is not clear whether the findings presented in this paper also apply to a memory task with visual input.

This study was supported in part by the Research Committee of the Graduate School of the University of Wisconsin with funds provided by the Wisconsin Alumni Research Foundation. Results of the study were presented at the Midwestern Psychological Association Meeting in May 1963.

REFERENCES

CALHOON, S. W. (1935). A comparison of the ability to reproduce one syllable words auditorially presented. *J. Exp. Psychol.* **18**, 621–32.

CONRAD, R. & HILLE, B. A. (1958). The decay theory of immediate memory and paced recall. *Canad. J. Psychol.* **12**, 1–6.

FRAISSE, P. (1945). Études sur la memoire immédiate. III. L'influence de la vitesse de présentation et de la place des éléments. La nature du présent psychologique. *Année psychol.* **35**, 29–42.

FRASER, D. C. (1958). Decay of immediate memory with age. *Nature, Lond.*, **182**, 1163.

MACKWORTH, J. F. (1962*a*). Presentation rate and immediate memory. *Canad. J. Psychol.* **16**, 42–7.

MACKWORTH, J. F. (1962*b*). The effect of the response upon the immediate memory span. *Canad. J. Psychol.* **16**, 120–7.

PEATMAN, J. G. & LOCKE, N. M. (1934). Studies in the methodology of the digit span test. *Arch. Psychol., N.Y.*, no. 167.

POLLACK, I. (1952). The assimilation of sequentially encoded information. Human Resources Lab., Memo Rt. No. 25, A.R.D.C. USAF.

POSNER, M. I. (1963). Immediate memory in sequential tasks. *Psychol. Bull.* **60**, 333–49.

(*Manuscript received* 3 *September* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 307–314
Printed in Great Britain

CHANGES IN HEART RATE AND PERFORMANCE AS A RESULT OF LOSS OF SLEEP

By D. W. J. CORCORAN

Medical Research Council, Applied Psychology Research Unit, Cambridge

Two theories of the effects of loss of sleep were compared against existing facts. The hypothesis that sleep deprivation reduces arousal was judged to be more adequate than the theory that arousal is increased by loss of sleep, the former failing only to account for raised physiological indices of arousal in some experiments. These are attributed to experimental conditions which compel the subject to make compensatory efforts to maintain adequate levels of performance. Under such conditions effort, not arousal, is reflected in the indices.

An experiment was conducted in which a constant high level of performance was not demanded of the subjects over a 60 hr. period without sleep. Heart rate and performance fell with loss of sleep.

I. INTRODUCTION

The present paper deals with a theoretical controversy which has built up in recent years in the field of sleep deprivation. Two apparently quite distinct and utterly opposed statements have been made concerning the neuro-physiological effects of loss of sleep. One (Wilkinson, 1958; Corcoran, 1963*b*) states that loss of sleep results in tendencies towards a lowering of 'arousal'; the other (Malmo & Surwillo, 1960) states that loss of sleep can, under certain conditions, result in heightened arousal.

Although these positions appear to be diametrically opposed, the opposition is partly semantic, resulting from different usages of the term 'arousal'. The present writer has used the term to refer to the *average* state of the organism over a relatively long period of time, operationally synonymous with 'probability of falling asleep' and perhaps neuro-physiologically related to degree of activation of the reticular formation or level of cortical facilitation over the long period. Malmo & Surwillo, on the other hand, appear to use the term in at least two ways: (*a*) in a similar way to the present writer, and (*b*) to refer to short-term states. Thus a subject who is deprived of sleep and taking short dozes would be said, according to the latter usage, to show periods of high arousal interspersed with periods of low. It is not altogether clear, however, whether Malmo & Surwillo hold that the periods between dozes are higher in arousal than the normal state, or whether the average level of arousal over the entire period is higher. Since the subject during this period is either asleep or in a state in which there is a high probability of his falling asleep, the present writer would disagree with either statement.

There are a number of experiments which demonstrate that if level of arousal is changed by deprivation of sleep, it must on average be lower. Wilkinson (1961*b*) showed that the level of performance of the subject deprived of sleep could be raised very nearly to control levels by giving him knowledge of his results. Corcoran (1963*b*) had subjects work to a required output rather than for a required time. It was found that performance under the former condition was superior to that in the latter condition after loss of sleep, but little different after normal sleep. Corcoran (1962) found that the decline in performance typical of subjects deprived of sleep could be reduced if they were exposed to loud white noise while the task was in progress. It was also shown (Corcoran, 1963*a*) that the day-to-day deterioration in a vigilance task during a 60 hr. vigil could be considerably decelerated by increasing the frequency at which signals were presented. In each of these experiments the results are compatible with the hypothesis that loss of sleep is associated with lowered arousal, but quite incompatible with the opposing explanation. Clearly, if it is because the subject is hyper-aroused that performance is low after loss of sleep without an incentive, in quiet, or with a low frequency of signalling, then the introduction of an incentive, noise or a high rate of signalling should lead to even more profound hyper-arousal and still poorer performance, rather than better performance.

Physiological evidence on the present issue is apparently equivocal. Although it is well established that the so-called 'physiological indicants of activation' (EEG, GSR, EMG, heart rate, temperature, respiration rate and blood pressure) show diurnal changes (Kleitman, 1939, 1950), attempts to relate changes in the indices to hours without sleep have not produced consistent results.

Tyler, Goodman & Rothman (1947) found, in subjects deprived of sleep for 100 hr., that EEG's taken from fronto-occipital regions showed a decrease in the percentage of time the records contained 8–12 cyc./sec. waves, and a corresponding increase in the time occupied by higher frequencies as the vigil proceeded. These changes occurred whether the subjects were resting or engaged in mental arithmetic. Superficially, this result indicates that level of activation was raised by loss of sleep. The authors, however, suggest that the sleepier the subjects became, the more effort was needed to remain awake. The EEG, they suggest, is sensitive to increased effort, and thus gave paradoxical results. Armington & Mitnick (1950) found changes in EEG indicative of lowered activation with loss of sleep, and Burch & Greiner (1958) found changes in the skin conductance in the same direction. Wilkinson (1961*a*) found that changes in muscle tonus were quite unrelated to hours without sleep. He also found that those subjects whose muscle tension had increased with loss of sleep showed the least decrement in performance at arithmetic.

Malmo & Surwillo (1960) used very nearly all the indices of arousal and found that they rose in all three of their subjects from the beginning to the end of a 60 hr. deprivation of sleep. The recordings were taken while subjects were engaged on a difficult tracking task in which poor performance was virtually prevented by having a coil, which was attached to the leg, heat to uncomfortable levels when performance become 'disproportionately low'. The performance of the three subjects in this experiment respectively dropped, rose and remained about the same, and it was further noted that the subject whose performance deteriorated showed the greatest *increase* in the physiological measures. It was on the basis of these results that the arousal theory arose. Sleep, it was stated, is a drive; thus loss of sleep is a drive deprivation. Like the deprivation of other drives, the prohibition of sleep causes increased levels of activation and performance therefore deteriorates owing to hyper-arousal. It was further suggested that the environment has to be sufficiently stimulating for loss of sleep to result in hyper-arousal, but no indication is given of how a sufficiently stimulating environment is defined. It was also not stated that low levels of arousal result after loss of sleep if the environment is insufficiently stimulating, but this may be what the authors imply. Indeed hyper-arousal is the only variable postulated to cause impaired performance as a result of loss of sleep. Thus one might reason that the more the 'stimulation' of a given situation is reduced, the less the situation will be effected by loss of sleep. Wilkinson (1961*b*) and Corcoran (1963*b*) have in fact shown that the more often a task is repeated, the more profoundly is it affected by loss of sleep. Repetition of a task is far more likely to reduce than increase its stimulation value, so that these results again support a de-arousal rather than an arousal theory.

Bunch, Cole & Frerichs (1937) and Bunch, Frerichs & Licklider (1938) have reported results which at first sight favour an arousal rather than a de-arousal theory. Rats deprived of sleep were shown to reach levels of performance above those of control animals. The sleep-deprived rats improved at learning water mazes as the time spent without sleep increased. After 24 hr. the improvement tended to cease, and after 48 hr. some decline in performance was observed.

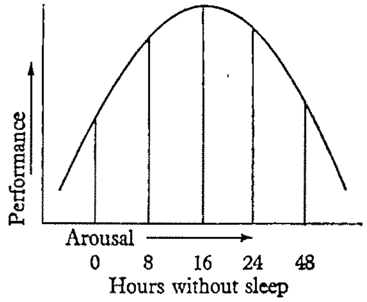


Fig. 1

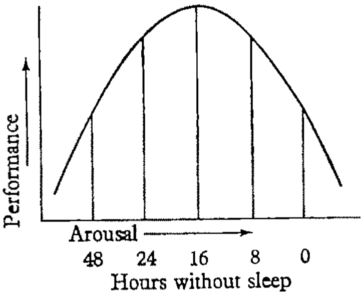


Fig. 2

Fig. 1. Explanation of the findings of Bunch *et al.* in terms of the arousal theory.

Fig. 2. Explanation of the findings in terms of the de-arousal theory.

These results can be explained on the theory that loss of sleep increases arousal, by the model shown in Fig. 1. This diagram illustrates the 'inverted-U relationship' postulated to exist between performance and arousal by such workers as Freeman (1940), Hebb (1955) and Stennett (1957). The control rats in the studies by Bunch *et al.* would, on this model, occupy a position at 0 in Fig. 1. With a progressive increase in time without sleep the level of arousal would increase and move through 8, 16, and 24 to 48 along the arousal continuum. Thus performance would improve up to 16 hr. without sleep and begin to decline only after 24 hr. deprivation, owing to excessively high levels of arousal.

An alternative explanation is that shown in Fig. 2. This suggests that under control conditions, the rats were already hyper-aroused, but that with loss of sleep arousal became progressively lower and performance thus improved by ascending the curve towards optimal levels through 8 and 16. After 24 hr. the level of arousal had begun to descend the curve and accordingly performance showed a small drop.

The question with which we should be concerned, therefore, is whether it is more reasonable to hold that the control animals in the Bunch *et al.* experiments were at point 0 in Fig. 1, or point 0 in Fig. 2. Since a water maze was used, it seems unlikely that the control subjects were low in arousal, and much more likely that they were excessively high in this respect.

A final study ostensibly in support of an arousal theory is that of Rust (1962). In this experiment it was shown that heart rate in rats increased with hours of sleep deprivation and it was therefore suggested that level of arousal had increased in the animals. Unfortunately it is a necessary experimental evil in work on loss of sleep that subjects who are deprived of sleep have also to be continually subject to stimulation, otherwise they fall asleep. The control rats in Rust's study were allowed 5–6 hr. of sleep per day over the 48 hr. deprivation whilst the experimentals were kept continually active on a mechanical activity wheel for this period, which means that as well as being without sleep the animals must have been near a state of physical exhaustion. Bunch *et al.* (1938) in fact found it necessary to discard some of their subjects after 48 hr. without sleep, because the animals were so physically exhausted as to be incapable of swimming through mazes, yet these animals were only handled to prevent them from sleeping. Since Rust's animals were kept continuously active on the activity wheel, it is hardly surprising that the heart rate of the experimental subjects increased and perhaps implausible to suggest that this increase was due to higher levels of arousal.

To summarize: the arousal theory of sleep deprivation finds itself in difficulty when it attempts to explain findings such as the effects of incentives, noise, signal frequency and the lowered physiological indices found by some investigators. The de-arousal theory, on the other hand, has difficulty in explaining the occurrence of raised indices of arousal, especially when these are accompanied by lowered performance. In order to get over this difficulty it is necessary to reintroduce the concept of effort, originally suggested by Tyler *et al.* Let us suppose that arousal fell in Malmo and Surwillo's subjects. The heated coil which was used virtually demanded that the subjects kept their performance at a fairly high level. As a result they had to put greater effort into remaining alert, otherwise their legs were burned. If the indices reflect effort as Tyler *et al.* and others (e.g. Farmer & Chambers, 1925) suggest, then the levels of the indices will increase with hours without sleep, since it will become increasingly difficult to stay awake *because arousal has been lowered.* The physiological indices of arousal may then become inversely related to level of arousal when sleepiness is extreme. The inconsistency of the results of past work may therefore be explainable on the hypothesis that varying amounts of effort were demanded of the subjects in the different experiments.

The present experiment is an attempt to control the factors which we suggest have led to paradoxical increases in the physiological indices, viz.

(1) The subjects were not kept active throughout a 3-day deprivation of sleep but were simply awakened as soon as they dozed.

(2) No method of inducing a constant high level of performance was used. In such a situation as this we would expect (*a*) that heart rate taken either when the subjects are resting or working would drop with loss of sleep, (*b*) that the decrease in heart rate from the beginning to the end of the deprivation would be accompanied by a decreased level of performance and (*c*) that during the period when the task is in progress the deterioration in performance would be mirrored by a lowering of heart rate. Malmo & Surwillo, on the other hand, would presumably expect (*a*) that resting heart rate would decrease with loss of sleep, (*b*) that working heart rate would increase, since the subjects were being stimulated, (*c*) for the drop in performance with

loss of sleep to relate to *increases* in heart rate, and (*d*) that a drop in performance from the beginning to the end of the task would relate to increases in the physiological index.

II. PROCEDURE

Subjects

Nineteen Naval volunteers between the ages of 17 and 21 served as subjects.

Heart rate

Heart rate was measured while a vigilance task lasting 30 min. was in progress. The subject lay on a couch in a warm room, stripped to the waist. After a preliminary cleansing of the skin with alcohol, electrode jelly was spread over the left pectoral region of the front of the chest and on the lateral surface of the wrist; suction electrodes were then applied. The electrodes were part of an electrocardiochronograph, an apparatus designed to record graphically each contraction of the left ventricle. Heart rate was sampled over pre-defined periods before and during the test. Resting heart rate was defined as the number of beats that occurred between 1½ and ½ min. before the test started. Working heart rate was assessed by means of four readings taken at the middle of the four quarters of the test, that is between 2½ and 3½ min. after the task started, 10½–11½ min. after the start, 17½–18½, and 25½–26½ min. after the start.

Task

The vigilance task used was a visual adaptation of the auditory task described by Bakan (1959). The subject lay with his head raised and electrodes in position. Above and behind his head was a projector which caused a series of single digits to appear on a screen facing him. The digits appeared to follow a random sequence, but at certain places within the sequence three odd digits were presented which were all different, e.g. 759. These were defined as 'signals' which subjects had to watch for, and report verbally by calling out the actual numbers making up the sequence. The digits were presented continuously for 30 min. either at a rate of 1 per sec. (the 'Slow' condition), or at 2 per sec. (the 'Fast' condition). The Slow condition contained 20 signals; the Fast condition contained 40 signals.

Method

The subjects were divided into four groups, of which two performed during a period of 60 hr. deprivation of sleep (NS), and two served as controls (C). The two NS groups, each consisting of 6 subjects, were treated as follows. During the first week of their six weeks' stay in the vicinity of the laboratory both groups were tested under the Slow condition once. On the subsequent week they slept normally on the Monday night, but did not sleep again until about 6.30 p.m. on the following Thursday. Between 7.30 a.m. and noon of Tuesday and Wednesday and between 1.30 p.m. and 6 p.m. on Thursday the subjects were tested (see Table 1). Six of the NS group were tested under slow conditions during the sleep deprivation and six were tested under fast conditions. In the slow group apparatus failure prevented the testing of two subjects on the third day of deprivation. The C groups were tested in exactly the same manner as the experimental, but slept normally during the intervening nights. Four of the C subjects received the slow task and three the fast.

III. RESULTS

Changes in heart rate during deprivation

Fast and slow groups were taken together for the following analysis. Table 1 shows the mean changes in heart rate and performance score in the C and NS groups on the practice session and on the subsequent three test sessions. Resting heart rate decreased between the practice session and day 1 to an almost equal extent in both groups. This probably indicates that subjects felt anxious about the task on the practice session, but were more relaxed on the second occasion. The reliability of the change in heart rate was tested using the binomial test. The decrease was found to be significant in the NS group ($P = 0{\cdot}006$) but not significant in the C group ($P = 0{\cdot}124$).

After a night without sleep (day 2) the heart rate of the NS group declined further but the decrease was not statistically significant. The heart rate of the C group also declined. Thus a single night without sleep did not decrease the heart rate to any greater extent than would be expected by simply repeating the test.

After two nights without sleep (day 3) the resting heart rates of the NS group dropped to a small and insignificant extent, but the rates of the C group tended to rise. This rise is probably due to the diurnal change from morning to afternoon. However, the change from day 2 to day 3 in the NS group was not significantly different from that of the C group.

Table 1. *Day-to-day trends in heart rate and performance*

Group	Practice, a.m.	Day 1, a.m.	Day 2, a.m.	Day 3, p.m.	*N*
NS					
Resting heart rate	77·3	70·6	66·2	64·3	12 (10 on day 3)
Working heart rate	79·1	71·9	65·3	59·5	
% signals detected	*95·5	79·1	37·5	16·0	
C					
Resting heart rate	77·1	69·9	66·1	74·4	7
Working heart rate	75·8	70·9	66·5	72·6	
% signals detected	94·3	86·7	86·8	84·7	

* On the practice day all subjects were tested under the slow (easier) conditions and the scores are therefore high.

Working heart rate (Table 1) showed no distinct differences from resting heart rate. Again, simple repetition of the task resulted in lower rates ($P = 0{\cdot}002$ in the NS group). A further significant drop occurred in the NS group from day 1 to day 2 ($P = 0{\cdot}038$); the decrease in the C group was not statistically significant. However, between day 2 and day 3 heart rate rose in the C group while it continued to fall in the NS group. The extent of the change is significantly different between the two groups; the Mann–Whitney test gave $U = 4, P = 0{\cdot}002$. This result indicates that in the NS group the raised heart rate to be expected in the afternoon was counteracted by loss of sleep. Thus loss of sleep tended to lower the heart rate.

To summarize these results:

(1) The heart rate recorded whilst the subjects were working showed no difference from that recorded when the subjects were resting.

(2) No difference in the heart rate was evident between sleepless and control subjects until the third day of the sleep deprivation. On this day the heart rates of control subjects increased, probably owing to the fact that the test took place in the afternoon, but the heart rates of the sleepless subjects continued to decline.

Relation of heart rate to performance

The differences between the effects of loss of sleep on the two conditions of the vigilance task are not relevant to this discussion and are discussed elsewhere (Corcoran, 1963*a*). Otherwise, in general from the beginning to the end of the run performance deteriorates and heart rate drops after loss of sleep. In order to assess whether the deterioration in performance was related to changes in heart rate the following procedures were applied. Heart rate and number of signals reported correctly were found for each half of the test on day 2 for each subject, and the value of the second half was subtracted from that of the first half. The differences for heart rate were then compared with the differences in performance; if both were positive or both were negative a plus was scored; if the direction of the change disagreed a minus was scored. If no relation existed between performance and heart rate during the task, there should on a null hypothesis be as many pluses as minuses. Ten of the twelve subjects showed plus values; this yields a two-tail probability equal to 0·038 on the binomial test. Thus the decrease in performance from the first to the second half of the test was related to a decrease in heart rate between halves. It cannot, however, be concluded from the analysis of the changes during the task that heart rate and level of performance were affected by a common causal factor, since performance is known to drop as a function of time on vigilance tasks, and heart rate is known to drop during rest.

It can be seen from Table 1 that the percentage of the signals detected on each day decreased profoundly in the NS group. All subjects in this group deteriorated both from day 1 to day 2 and from day 1 to day 3 ($P < 0{\cdot}01$ for both comparisons). In the C group there were no reliable changes in performance from day to day.

Since heart rate was shown to decrease with loss of sleep, changes in heart rate and performance would seem to coincide in the NS group. However, although performance was poor in the NS group on day 2, heart rate was no lower in this group than in the C group, where performance was unchanged from the previous day. Thus we cannot conclude that heart rate determined level of performance, but rather that loss of sleep affected both heart rate and performance; both measures perhaps reflecting level of arousal.

IV. DISCUSSION

To the extent that heart rate is a valid measure of level of arousal, the results of the present experiment indicate that loss of sleep was associated with a drop in arousal. It must be emphasized, however, that the conditions under which subjects were tested were carefully designed to be unmotivating, and to involve no physical activity. Under such conditions results consistent with the predictions were obtained.

If the concept of arousal is to be used at all meaningfully, it must refer to the state of the organism. It must be kept quite distinct from the physiological mechanisms used in measuring it, since they can be affected by incidental experimental conditions, and thus give equivocal results.

Holding to the hypothesis that loss of sleep lowers arousal makes it necessary to introduce, as Tyler *et al.* (1947) did, the concept of 'effort'. 'Effort' and 'arousal' are quite distinct. While the latter refers to a state of the organism, the former refers to an activity; indeed the effort to remain awake is frequently increased because arousal has become decreased. It is therefore suggested that arousal falls with loss of sleep, but when this occurs the experimental conditions may demand compensatory effort. The effort, not the lowered level of arousal, is reflected in the indices.

The author is indebted to Mr D. E. Broadbent and Dr R. Conrad for their advice and to the Royal Navy for supplying subjects, equipment and assistants.

REFERENCES

Armington, J. C. & Mitnick, L. L. (1950). EEG and sleep deprivation. *J. Appl. Physiol.* **14**, 247–50.

Bakan, P. (1959). Extraversion-introversion and improvement in an auditory vigilance task. *Brit. J. Psychol.* **50**, 325–32.

Bunch, M. E., Cole, A. & Frerichs, J. (1937). The influence of 24 hours of wakefulness upon the learning and retention of a maze problem in white rats. *J. Comp. Psychol.* **23**, 1–11.

Bunch, M. E., Frerichs, J. & Licklider, J. C. R. (1938). An experimental study of maze learning ability after varying periods of wakefulness. *J. Comp. Psychol.* **26**, 499–514.

Burch, N. R. & Greiner, T. H. (1958). Drugs and human fatigue: GSR parameters. *J. Psychol.* **45**, 3–10.

Corcoran, D. W. J. (1962). Noise and loss of sleep. *Quart. J. Exp. Psychol.* **14**, 178–82.

Corcoran, D. W. J. (1963*a*). Doubling the rate of signal presentation in a vigilance task during sleep deprivation. *J. Appl. Psychol.* **47**, 412–15.

Corcoran, D. W. J. (1963*b*). Individual differences in performance after loss of sleep. Unpublished Ph.D. Thesis, University of Cambridge.

Farmer, E. & Chambers, E. G. (1925). Concerning the use of the psycho-galvanic reflex in psychological experiments. *Brit. J. Psychol.* **15**, 237–54.

Freeman, G. L. (1940). The relation between performance level and bodily activity level. *J. Exp. Psychol.* **26**, 602–8.

Hebb, D. O. (1955). Drives and the C.N.S. (Conceptual Nervous System). *Psychol. Rev.* **62**, 243–54.

Kleitman, N. (1939). *Sleep and Wakefulness.* Chicago: University of Chicago Press.

Kleitman, N. (1950). The sleep-wakefulness cycle. In H. Abramson (ed.): *Problem of Consciousness.* New York: Josiah Macey, jr. Foundation.

Malmo, R. B. & Surwillo, W. W. (1960). Sleep deprivation: changes in performance and physiological indicants of activation. *Psychol. Monogr.* **47**, no. 15, 1–24.

Rust, L. D. (1962). Changes in bar pressing performance and heart rate in sleep deprived rats. *J. Comp. Physiol. Psychol.* **55**, 621–5.

Stennett, R. G. (1957). Performance level and level of arousal. *J. Exp. Psychol.* **54**, 54–61.

Tyler, D. B., Goodman, J. & Rothman, T. (1947). The effects of experimental insomnia on the rate of potential changes in the brain. *Amer. J. Physiol.* **149**,, 185–93.

Wilkinson, R. T. (1958). The effects of lack of sleep on perception and skill. Unpublished Ph.D. Thesis, University of Cambridge.

Wilkinson, R. T. (1961*a*). Effects of sleep deprivation on performance and muscle tension. In *CIBA Symposium*, ed. by G. E. W. Wolstenholme and M. O'Connor. London.

Wilkinson, R. T. (1961*b*). Interaction of lack of sleep with knowledge of results, repeated testing and individual differences. *J. Exp. Psychol.* **62**, 263–71.

(*Manuscript received* 18 *April* 1963; *revised* 24 *August* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 315–320
Printed in Great Britain

PERCEPTION OF VERTICALITY IN ADULT LIFE

By ANN D. M. DAVIES
Medical Research Council Unit for Research on Occupational Aspects of Ageing, Liverpool

AND G. W. H. LEYTHAM
Department of Psychology, University of Liverpool

Comalli, Wapner & Werner's (1959) developmental study on the perception of verticality has been repeated and extended using eight men and eight women matched for distribution of intelligence in each of the age decades from 20–29 to 70–79. Following Comalli *et al.* subjects were seated in a dark room in the upright position or tilted at ±30°. They were asked to adjust a luminescent rod set initially at ±10°, ±30°, until they thought it was upright. Estimates of verticality were found to depend on the starting position of the rod and to be related to non-verbal intelligence.

The subjects, unlike those of Comalli *et al.*, located the vertical to the side opposite to the body tilt throughout the 20–79 age range.

I. INTRODUCTION

Comalli *et al.* (1959) describe a study which suggests that there are progressive changes with age in response to body tilt. They found that if young or middle-aged male subjects were tilted to the left, a luminescent rod in a dark room was perceived as being vertical when it was in fact tilted to the right of plumb. However, a group of fifteen elderly subjects were found to perceive the apparent vertical as being tilted to the same side as their body tilt, that is, when they were tilted to the left they thought that the rod was vertical when it was, in fact, tilted to the left. Although Comalli *et al.* report that they found 'striking differences in effect of body tilt depending on age level', the differences they demonstrated were very small. When subjects in the old group were tilted 30° to the left, their mean position of apparent vertical was only 0·4° left. Comalli *et al.* unfortunately did not investigate subjects between the ages of 50 and 65, but suggest by interpolation that this is the range in which the apparent reversal takes place. Although the authors give no information as to the basis of selection of subjects, there is some evidence to suppose that the groups may have been of different intelligence. The young and middle-aged subjects were from a University Evening College, whereas the old subjects were drawn from a community old-age club.

The present study is an attempt to verify the developmental trend in perception of verticality found by the previous authors, using both men and women throughout the age range 20–79, matching decade groups for vocabulary level, and controlling the level of non-verbal intelligence.

II. PROCEDURE

The procedure was essentially the same as that of Comalli *et al.*, in order that direct comparison might be made with their data.

The subject was blindfolded outside the laboratory and did not see the experimental situation in the light. He or she was led into a darkened room and seated in a chair which could be placed in one of three positions, 30° left, 30° right, or in the upright position. The chair had a headrest and firm sides and back, so that the subject's body was held in place securely. The chair was 6 ft. in front of a luminescent rod, pivoted at its centre. The subject's task was to instruct the experimenter to adjust the rod until it appeared vertical. This task was carried out under four conditions of rod starting positions (SP) (30° left, 10° left, 10° right and 30° right), and the three conditions of body tilt (BT) mentioned above. There were thus twelve experimental conditions in all.

A different random order of trial presentation was used for each subject. After the experiment the room lights were switched on and the subject was asked to set the rod to vertical under each position of tilt. This was to ascertain that the subject had understood the nature of the experiment. (Two subjects who had misunderstood it were replaced.) In every instance the subject was blindfolded after making each adjustment until the beginning of the next trial.

Some comments on the procedure

The experiment of Comalli *et al.* and thus, of necessity, the present experiment share two methodological flaws. First, there is no experimental condition in which the rod is upright in its starting position. Secondly, as Mulholland (1962) has pointed out, in experiments where the subject merely says 'stop' when the rod appears vertical, there is no exploration of the range of positions called 'vertical'. This method uses the mean of a lower limen (obtained only in an ascending series) and an upper limen (obtained only in a descending series). This estimate of the point of subjective equality is not necessarily the same as that obtained by taking the mid-point of the range of positions called 'vertical' in a complete series.

One might add that, as the subject proceeded with the next trial as soon as he made his judgement, a further confounding element is introduced. The judgements were not made in ideal conditions where the subject had had a chance to become physiologically adapted to the position of tilt. It is possible that the error made by a subject on any trial depends not only on the body tilt and starting position for that trial, but also on the direction of tilt on previous trials.

III. SUBJECTS

The choice of subjects is described in more detail than is customary: first, because the study of Comalli *et al.* used an incomplete age range, and secondly because it was possible in the present work to maximize the range of intelligence *within* each group whilst keeping intellectual differences *between* groups as small as possible. Ninety-six subjects were used: eight men and eight women in each of the age decades from 20–29 to 70–79. The subjects were drawn from a pool of about 700 subjects in the Medical Research Council's Merseyside Voluntary Research Panel, each of whom had completed the Raven's Progressive Matrices test (1938 version) testing non-verbal intelligence, together with the Mill Hill Vocabulary Scale (Synonyms Senior B). Subjects were chosen primarily on the basis of their vocabulary scores, because cross-sectional studies indicate consistently that vocabulary does not decline with age. The highest and lowest vocabulary scores common to all age-groups were 40 and 24. Subjects were therefore chosen with scores as near as possible to 40, 34, 29 and 24.

Additionally, in order that non-verbal intelligence might be controlled, the Progressive Matrices scores of subjects in each vocabulary bracket were chosen so that they fell steadily with age. This is also in accordance with the results of cross-sectional studies, which show that scores in this test decline with age.

IV. RESULTS

For a preliminary analysis, the errors made by a subject were represented graphically. Angles to the right and left of plumb were designated positive and negative respectively. The errors made by men and women were plotted separately as a function of age, for each of the 12 experimental conditions, e.g. BT = 0°, SP = −10°. Although individual scatter diagrams could be taken to suggest some slight developmental variation of error with age, some positive and some negative, the diagrams as a whole suggested no developmental trend. To confirm this conclusion the distributions were examined numerically to see whether or not the gradients (i.e. the developmental trend with age) of each distribution were significantly different from a random variation about zero slope. The least squares method was used to calculate the value of the gradient for each experimental condition and the significance of the value found was examined using Student's *t* test.

Table 1. *Gradient of slope showing developmental trend* (*in degrees per year*) *in each experimental condition*

		BT (men)			BT (women)		
SP		−30°	0°	+30°	−30°	0°	+30°
+30°	Pure	+0·032	+0·027	+0·094	+0·068	+0·132**	+0·142*
	All	+0·122	+0·103*	+0·221*	+0·028	+0·066	+0·057
+10°	Pure	+0·057	−0·003	+0·217**	−0·029	+0·068	+0·144
	All	+0·078*	+0·027	+0·212**	−0·056	+0·018	+0·100
−10°	Pure	−0·104*	−0·031	+0·081	−0·098*	+0·038	+0·022
	All	−0·133*	+0·052	+0·118	−0·101	+0·056	+0·097
−30°	Pure	−0·045	−0·011	+0·060	−0·011	+0·054	−0·051
	All	−0·065	+0·032***	+0·060	+0·061	+0·053	+0·047

* $P < 0{\cdot}05$; **$P < 0{\cdot}01$; *** $P < 0{\cdot}001$.

During testing a few subjects had given responses which suggested that they were completely disorientated. Two of them, for instance (one man and one woman, both in their seventies) claimed that they saw the rod on the ceiling and felt that they were looking up at it. Since extreme errors affect disproportionately the value of a gradient computed by the least squares method, the gradient was calculated twice for each experimental condition. In the first case, the few extreme errors (those greater than about three standard deviations from the mean) were excluded (Pure). The second computation included all possible data (All). The results for both men and women are presented in Table 1. It will be seen that all the values of the gradients are very small and the greater number of them statistically insignificant. The results for men show no great consistency with those for women; furthermore, no pattern or trend in either the direction or the size of the gradient is evident.

Comalli *et al.* confirmed their result by carrying out an analysis of variance. However, there are difficulties inherent in using this method with the data obtained from an experiment of this sort: the variation in the estimation of verticality is relatively slight in conditions where the subject sits upright; there is much greater variability when the subject is tilted to the left or to the right. Moreover, there are large and small scatters associated with the large and small values of starting position. It is thus doubtful whether the assumption of homogeneity of variance is satisfied by the data resulting from experiments of this kind. However, so that close comparison between the present study and that of Comalli *et al.* might be possible, an analysis of variance for men and for women was carried out (excluding a small quantity of extreme data as described below).

The analysis of variance summarized in Table 2 is to be compared directly with table 2 of the Comalli paper. It will be seen that in the present work the interaction A × BT (age by body tilt) is not significant ($P > 0{\cdot}05$) either for men or for women.

Table 2. *Development and perception of verticality*

		Men			Women		
Source of variation	D.F.	Sums of squares	*F*	*P*	Sums of squares	*F*	*P*
Between subjects							
Between age levels (A)	5	671	1·55*	> 0·05	490	1·50*	> 0·05
Between subjects within age levels	38	3287	5·81†	< 0·001	2472	4·07†	< 0·001
Within subjects							
Between body tilt (BT)	2	3301	29·4‡	—	1095	8·26‡	—
Between starting position (SP)	3	3207	37·7§	< 0·001	5901	5·11§	< 0·01
BT × SP	6	628	7·03†	< 0·001	817	8·51†	< 0·001
A × BT	10	907	1·61‡	> 0·05	793	1·20‡	> 0·05
A × SP	15	412	0·97§	> 0·05	741	1·28§	> 0·05
A × BT × SP	30	475	1·06†	> 0·05	660	1·38†	> 0·05
Pooled S × BT	76	4274	3·77†	< 0·001	5036	4·14†	< 0·001
Pooled S × SP	114	3230	1·90†	< 0·001	4390	2·41†	< 0·001
Pooled S × SP × BT	228	3406	—	—	3651	—	—
Total	527	23798	—	—	26046	—	—

* Tested against between S within age-groups.
† Tested against pooled S × SP × BT.
‡ Tested against pooled S × BT.
§ Tested against pooled S × SP.

Additional findings

While the main aim of the investigation was to test the conclusion of Comalli *et al.*, the data which were obtained enabled us to draw some further conclusions. When the performance of the *total* male and the *total* female population was analysed it was clear that the error in estimating verticality does depend markedly on rod starting position and on body tilt. When the chair is tilted to the left errors made are mostly to the right (and vice versa). When the chair is upright an error of the same sign as the rod starting position is made. This starting position effect is well known in many psychophysical demonstrations. In fact the present data are consistent with an error of about 0·1° per degree SP for women and a little less for men.

The data were next analysed to see whether there was any relationship between magnitude of error and intelligence level. The total error made by each subject on the twelve trials was calculated (irrespective of sign). A frequency distribution of errors was made and three men and three women with extremely large total errors (130° +) were discarded for this correlational analysis and for the analysis of variance described above. These six formed a tail separated from the rest of the distribution and it was felt that these extreme results might give spurious significances. (No scores could be obtained from the two subjects who saw the rod on the ceiling.) The errors made by each subject in each of the three conditions of body tilt taken separately were also obtained. Product moment correlations were computed between these scores and those on the Progressive Matrices test, and between all of these scores and age. The results are shown in Table 3.

Table 3. *Correlation between error, age and non-verbal intelligence*

		Age	PM
Progressive Matrices (PM)		−0·62***	
Total errors		+0·12	−0·21*
Error made at	+30°	+0·09	−0·19
body tilt:	0°	+0·22*	−0·34***
	−30°	+0·06	−0·05

Correlation between PM and error holding age constant = −0·27** (BT = 0°).
Correlation between age and error holding PM constant = +0·01 (BT = 0°).

* $P < 0\cdot05$; ** $P < 0\cdot01$; *** $P < 0\cdot001$.

It can be seen from this table that when the subject is in the upright position, error correlates positively with age ($P < 0\cdot05$) and negatively with non-verbal intelligence ($P < 0\cdot001$). However, non-verbal intelligence declines with age ($P < 0\cdot001$). When PM score is held constant statistically the correlation between age and error no longer reaches significance ($r = +0\cdot01$, $P > 0\cdot05$). When age is held constant there is still a significant correlation between PM score and error ($r = -0\cdot27$, $P < 0\cdot01$). There are no significant correlations between error and any of these variables when the angle of body tilt is ± 30°.

V. DISCUSSION

The present study fails on any analysis to substantiate the findings of Comalli *et al.* that older subjects locate the vertical to the same side as body tilt. Their conclusion that subjects in old age tend to revert to a more ego-centric organization of space cannot therefore be regarded as proven, insofar as it is based on this evidence.

The present experiment in fact finds that for all age decades in the adult life span, errors of opposite sign to body tilt and of the same sign as starting position are made. As might be expected, the error appears to be very roughly proportional to starting position. It was found, too, that the error obtained by averaging over all conditions (BT, SP and age), was not zero but a little less than + 1°. One cannot exclude the possibility that this appeared because of some subtle asymmetry of the room

(Rosman, 1960); it seems much more probably associated with the fact that the population was predominantly right-handed (91 out of 96).

In the upright position, the error the subject makes is smaller the greater his non-verbal intelligence. (This intellectual advantage disappears when the subject is tilted.) The error also increases very slightly with age. However, non-verbal intelligence declines with age and an analysis suggests that in an experiment using a population of all ages, but of uniform non-verbal intelligence, no significant variation of error with age would be found.

The authors would like to acknowledge the help and advice given by Dr Alastair Heron, Director of the Medical Research Council's Unit for Research on Occupational Aspects of Ageing, and by Dr M. G. Davies, Department of Physics, University of Liverpool. They would also like to thank Mrs B. G. Brechling for help in obtaining the subjects.

REFERENCES

COMALLI, P. E. Jr., WAPNER, S. & WERNER, H. (1959). Perception of verticality in middle and old age. *J. Psychol.* **47**, 259–66.

MULHOLLAND, T. B. (1962). Method of limits and implicit standards: a comment on two experiments. *Percept. Mot. Skills*, **14**, 283–90.

ROSMAN, R. R. (1960). The tilting-chair tilting-rod task: a methodological note. *Percept. Mot. Skills*, **10**, 9–10.

(*Manuscript received* 17 *July* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 321–331
Printed in Great Britain

AN ANALYSIS OF THE PURSUIT ROTOR LEARNING OF CHRONIC PSYCHOTICS*

By P. L. BROADHURST† and ANNE BROADHURST†
Institute of Psychiatry, University of London

A sample of forty-two chronic schizophrenics were randomly assigned to one of two groups given three sessions of either massed or spaced practice on a pursuit rotor, and a group of sixteen depressives were given massed practice only. The results suggest that there are differences between the learning curves of these groups when compared with those of normals. Notable among these differences is the absence of reminiscence after massed practice, and the failure of spaced practice to give the usual distinct superiority over massed. Statistical analysis of the data confirms the inferiority of the psychotics in respect of level of learning, but fails to show any other marked differences in the nature of the learning curves. The results are discussed primarily in the context of Hullian learning theory.

The experimental investigation of slowness of mental functioning among psychotics has had a long history (e.g. Boring, 1913), and it is of interest in the light of what follows that Kraepelin (1913) even attempted the measurement of reminiscence. Our own approach to this problem developed from work by one of us (Broadhurst, 1958*a*, *b*) which showed that on certain learning tasks schizophrenics, given sufficient practice, could achieve speeds of performance displayed by normal subjects before practice. To account for these findings it was postulated that schizophrenics suffer from an excess of reactive inhibition (Hull, 1943), because there is evidence in the literature (reviewed by Broadhurst, 1958*b*) that the inferiority of schizophrenics is accentuated when they are given massed, as opposed to spaced, practice—the former favouring the development of reactive inhibition. It was therefore decided to explore this hypothesis by using massed and spaced practice on a pursuit rotor task. This device has been intensively used in this country and elsewhere (Bilodeau & Bilodeau, 1961) to test predictions derived from Hullian learning theory, so that adequate norms for non-psychotic populations are consequently available. It also has the advantage of having already been shown to be suitable for use with psychotics (Huston & Shakow, 1948).

METHOD

Apparatus

The apparatus and procedure were the same as used by Eysenck (1956*b*) in testing a group of male university students, fifty of whom were given massed practice and twenty-five spaced. The pursuit rotor used has been described (Eysenck, 1956*a*) and illustrated (Eysenck, 1957, figs. 7 and 8, facing p. 80) elsewhere. Briefly, it consists of a gramophone-like turntable, 10 in. in diameter, which has a metal spot $\frac{7}{10}$ in. in diameter embedded into its surface, the centre of which is 3¼ in. from the centre of the turntable. The subject's task is to maintain contact between this spot and the end of an articulated stylus, 10½ in. long. Such contact activates timers which record the length of time of the contact in successive periods of 10 sec. The turntable revolved in a clockwise direction at a speed of 72 rev./min., and was mounted at a height of 36½ in. from the floor on which the subject stood. The illumination at the turntable was measured at 35 foot-candles with an Avo no. 2 light meter.

---

* This study is dedicated on publication to the memory of Donald Walton. A preliminary account of the findings has been presented to the British Psychological Society (Broadhurst & Broadhurst, 1959).

† Now at the University of Birmingham.

Subjects

Two distinct diagnostic groups of male psychiatric patients were tested at Springfield Hospital, London, S.W. 7. There were sixty-seven schizophrenics and thirty depressives, a total of ninety-seven. All the patients showed marked psychotic features, and an endeavour was made to secure only those for whom the differential diagnosis was not in question. To this end, four subjects were later deleted from the depressive group on the grounds of doubtful diagnosis. As will be seen from Table 1, which gives the age and length of stay in hospital of all subjects whose test scores were eventually included in the analyses (see 'Results' section below), our subjects may be regarded as chronic patients, especially when it is considered that in many cases this was not the first stay in hospital. Thus, 31 % of the schizophrenics and 38 % of the depressives shown in Table 1 had been hospitalized at least once before. It was not possible to ensure the inclusion of only those subjects who had not undergone some form of physical treatment (either electroconvulsive or insulin therapy), or who were not currently receiving some form of medication—indeed, only 14 % of the schizophrenics and 13 % of the depressives shown in Table 1 fell outside one or other of these categories—but one schizophrenic who had been leucotomized was excluded from the groups as finally composed (see below). The schizophrenic patients were allocated strictly at random to a massed or a spaced practice group. There were insufficient depressives available to dichotomize this group in the same way, and all were consequently given massed practice. The two schizophrenic groups thus formed are very comparable in respect of age, but the depressives are a substantially older group, as might be expected.

Table 1. *Subjects*

			Age at testing (years)		Length of stay in hospital (years)	
Diagnosis	Practice group	*n*	*M*	S.D.	*M*	S.D.
Schizophrenic	Spaced	21	38·5	9·0	6·8	6·0
	Massed	21	36·9	8·5	5·1	6·6
Depressives	Massed	16	58·5	9·5	1·2	1·5
Total		58				

Procedure

The instructions given the subjects were as follows: 'First I'll show you what I want you to do, and then I'll tell you how I want you to do it. Which hand do you use for writing? For this you will use some of the same muscles as for writing, so I want you to hold this pointer in your right hand.' (All subjects said they were right handed, and in cases of doubt this was checked by asking the patient to write his name.) 'Now you see this thing like a gramophone and you see this spot moving round. Stand comfortably in front of it here, and, holding the pointer lightly and not pressing down hard with it, move your hand and arm round so as to keep the pointer on this spot as much as possible.' (Experimenter demonstrates.) 'Don't press too hard on the spot—just keep the tip of the pointer on it lightly and as much as possible, always following it round and round. I want you to keep your eyes on the spot—don't look at other things or you'll find it harder to keep the pointer on it. Do you understand?' (Spaced practice group only.) 'Now that's what I want you to do, but you'll have lots of rests. I will say "Stop", and you must stop *immediately*, put the pointer down here, and then sit down here. Then I'll say "Get ready to start again"—you will get up, and pick up the pointer. I'll say "Are you ready? Start!" and you must start immediately. Is that clear ?' (Massed practice group only.) 'Now that's what I want you to do, and I want you to practise that for a time, starting immediately I say "Are you ready? Start!" Then we will have a rest before going on again. During each rest, stop immediately I say "Stop", put the pointer down here, and then sit down here. Do you understand?' Any phrases of these instructions were repeated as necessary, or in answer to the subject's questions, and repetition was made at the end of the 10 min. break between the sessions.

There were three sessions of practice of 5 min. each, separated by the two rest pauses of 10 min. duration. The massed practice group practised continuously for the 5 min.: for the spaced practice group a rest of 30 sec. was allowed between each 12½ sec. period of performance. The first 2½ sec. of each such period were unscored, in order to allow the subject time to get the pointer on to the turntable at the beginning of each trial, but the number of trials was reduced to 24—as opposed to the equivalent 30—in order to equate this group with the massed practice group for the total amount of time spent in practice. The only departure from this procedure occurred at the beginning of Session I, when, following Eysenck, the spaced practice group was given three consecutive 10 sec. trials without a break, i.e. 30 sec. practice. The last of these three trials counted as the first of the session, so that an additional 20 sec. of practice was given in this session. The purpose of this arrangement was to secure an adequate block of practice for comparing the three psychotic groups with respect to their initial ability on the test, and, in particular, to check that the random assignment of the schizophrenic subjects to the massed and spaced groups had not inadvertently resulted in any difference between them in this respect. Anticipating the results a little allows us to note that an analysis of variance of the total time on target in the first 30 sec. for three psychotic groups yielded an insignificant F ratio between groups, confirming the absence of any differences in initial ability.

The attempt to administer the test had to be abandoned in the case of twenty-six patients (five depressives and twenty-one schizophrenics) usually because of failure to co-operate. The co-operation secured from the rest of these psychotic subjects was in general good, but it was found that the instructions given at the beginning of the test were not always sufficient, and that, in order to ensure continued performance, it was necessary to repeat appropriate parts of the instructions in an encouraging manner, especially during the first session of practice.

No particular activity was specified for the rest intervals, there being evidence that variations in pre-practice activities have little effect on pursuit-rotor learning (Ammons, 1951). The experimenter engaged in casual conversation during the longer (10 min.) rest intervals: of the shorter (30 sec.) rests allowed the spaced practice group, much of the time—an average of 15 sec. —was occupied in sitting down and getting up again.

RESULTS

As is sometimes the case, even with normal subjects practising the pursuit rotor (e.g. Irion, 1949) there were some subjects who, while co-operative, and despite the repetition of the instructions, still did extremely poorly throughout the whole test, and showed no evidence of learning. It therefore became necessary to adopt the procedure usual in such cases, and to apply some minimal criterion for inclusion of a subject's score in the final group. The criterion arbitrarily chosen was that the subject must show at least one 10 sec. period in which he maintained contact with the spot for 15 % of the time (1·5 sec.), after deducting his performance during the very first period. Application of this criterion excluded a further eight of those tested, leaving us with the final total of fifty-eight, divided as shown in Table 1. Of these, twenty-seven (sixteen schizophrenics and eleven depressives) had received encouragement as described above.

The results for the three groups are shown in Fig. 1. For the purpose of comparison the results for the normal groups tested by Eysenck (1956*a*, *b*) in exactly similar fashion are included in the figure.

Striking differences between the curves for normal and psychotic groups are seen, but before proceeding to consider them, it is important to enquire into the extent to which the former can be used as controls for the latter. One doubtful feature is the age difference, since it was not possible to match our psychotic groups for age with Eysenck's student subjects, few of whom were over 25 (Eysenck, 1956*a*). As Table 1 shows, our subjects were much older. Lacking definitive data on the effect of age on pursuit rotor learning, though Huston & Shakow (1948) do report the absence of correlation among schizophrenics, we can only investigate the effect of age within our groups, a procedure which is admittedly not completely satisfactory. However, the correlations between age and the total time on target in each of the three sessions for each of the three psychotic groups—nine altogether—show only one to be significant at the 5 % level, and that in the direction suggestive of a *better* performance with increasing age. As far as they go, therefore, these results discount age as a major determinant of the differences observed. Eysenck (1956*a*) reaches a similar conclusion regarding age differences within his student group given massed practice, though it should be noted that Griew & Lynn (1960; Lynn, 1962) suggest that reactive inhibition increases with age. The other consideration affecting our use of Eysenck's data for control purposes is intelligence level. Measures of I.Q. were not available for

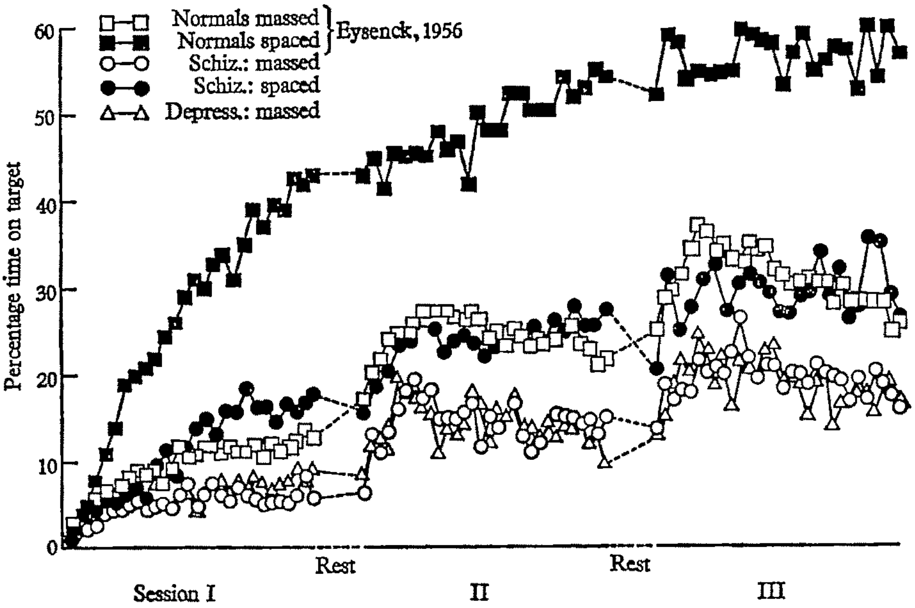


Fig. 1. The figures show the work curves—plotted from the mean scores for successive 10 sec. periods—for the three psychotic groups (the three bottom curves at the start of Session II). The curves for the normal groups from Eysenck's studies (the top two curves at the start of Session III) which serve as controls are included for comparison.

all psychotic patients, but if those that were are representative the mean would be lower than the 120 which Eysenck estimates as the mean for his group. However, it seems likely that intelligence is a relatively unimportant factor in pursuit-rotor learning (Huston & Shakow, 1948), and that the exclusion of many subjects incapable of learning the task for one reason or another would further diminish its importance. Nevertheless, as with age, we are not able definitely to exclude any possible influence, and while these deficiencies seem to us unimportant, they should be borne in mind.

Inspection of Fig. 1 shows several noteworthy features. The first is the generally lower level of the performance of the psychotic groups. This is doubtless an example of the psychomotor slowness frequently associated with the performance of psychotics. The second striking feature is the apparent absence of reminiscence among the massed practice groups of psychotics, which is in strong contrast with the findings for the normal group. For them reminiscence, which is measured (Eysenck, 1956*a*, *b*) as the increase in performance of the first post-rest 10 sec. period over the average of that on the last three pre-rest periods, is highly significant in both cases, that is, after both rest periods. The comparable psychotic scores for massed practice show no such trend in any of the four cases calculated in this way and presented at the left of Table 2. Thus, in two cases there is a rise in the mean score, but in the other two a fall. Moreover, none of these differences is significant when evaluated by means of Sandler's (1955) *A* test for correlated differences. However, both the psychotic groups under massed practice show the same 'warm-up' or extinction increment effect in the first minute of resumed practice after the rest period, as do the normals (see Fig. 1). This rise in the curve Eysenck (1956*b*) interprets as due to the extinction of conditioned inhibition, arising from the absence at first of the normal build-up of reactive inhibition during the resumption of practice. Conditioned inhibition, developed as a result of the reinforcement provided by the involuntary rest pauses occurring when the negative drive (reactive inhibition) surpasses the positive drive or motivation for the task, ceases to be reinforced in this way and consequently dissipates, as a normal extinction effect. Rid of this negative component, the scores show an increase until, with the passage of time, reactive inhibition is built up again, causing the decrement seen. Eysenck & Maxwell (1961) give a succinct statement of the theory as it relates to pursuit rotor learning, and Jensen (1961) and Adams (1961) critical analyses of it; but see also Feldman (1963).

Table 2. *Significance of differences pre- and post-rest: practice-groups*

(Units are percentage of time on target.)

		Massed					
		Depressives		Schizophrenics		Spaced schizophrenics	
Session		*M*	S.D.	*M*	S.D.	*M*	S.D.
I–II	Pre-rest	8·3	8·8	6·3	7·1	17·0	11·3
	Post-rest	7·9	7·8	8·2	7·4	13·5	9·6
	Significance	N.S.		N.S.		$P < 0{\cdot}02$	
II–III	Pre-rest	11·2	10·8	14·2	8·8	26·0	16·9
	Post-rest	12·6	8·7	13·6	6·8	20·4	15·5
	Significance	N.S.		N.S.		N.S.	

Turning to the spaced practice groups, we find that the generally beneficial effect of spacing practice is less among psychotics than among normals. The normals given spaced practice reach levels between two and three times as high as those given massed, whereas with schizophrenics the spaced performance is rarely even twice as good as the massed. No reminiscence effect is found in either spaced group, and this absence in the normals is interpreted as due to the interpolation of the 30 sec. rest periods during the practice sessions. These allow the dissipation of reactive inhibition as soon as it forms instead of allowing it to accumulate, and dissipate during the 10 min. rest interval between the sessions. But among the schizophrenics, this rest period has apparently depressed performance upon resumption of practice, and this difference between pre- and post-rest performance, evaluated as before, may be seen (at the right of Table 2), to have reached a significant level in respect of the first 10 min.

rest period. A third feature of the schizophrenics' spaced practice curve is the apparent warm-up effect it shows, which is not normally seen in spaced practice curves (cf. the normals at the top of Fig. 1). The effect is marked at the beginning of Session II, though the resemblance to the curve for the normals under massed practice at this point is almost certainly fortuitous.

Table 3. *Summary of results of analysis of variance of the transformed scores: massed practice*

	Source	D.F.	M.S.	F	Significance
	Between groups (G)	2	109·08	16·8	*
	Between sessions (S)	2	440·83	67·8	***
	G × S	4	6·50	54·2	***
	Within groups (Individuals)†	249	4·90	40·8	***
	Residual including within sessions	7482	0·12	—	—
Within sessions	Periods (P)	87	1·04	9·45	***
	G × P	174	0·08	0·08	N.S.
	Residual	7221	0·11	—	—
	Total	7739	—	—	—

† Including individuals × sessions interaction. N.S. = not significant; * = significant at the 5 % level. ** = significant at the 1 % level. *** = significant at the 0·1 % level.

In order to assess the nature of the overall trends shown in Fig. 1 somewhat more rigorously, further analyses were performed. In the first, the three groups given massed practice, that is, normals, depressives and schizophrenics, were compared by means of the analysis of variance. A previous analysis of the normals only under massed practice, reported in detail elsewhere (Eysenck, 1956*b*), has shown the need for transforming the data to achieve homogeneity of variance, and the same square root transformation was therefore applied to the psychotic data. The analysis referred to had, in brief, shown: (1) that the normals' mean score varied significantly between sessions and between 10 sec. periods within sessions, (2) that the differences could be expressed in terms of (*a*) an average level for the whole session, (*b*) a linear component of change (slope of the curve), and (*c*) a curvature component; and that (3) whereas the average level and slope components varied significantly for the three sessions, the curvature component did not. The results of the present analysis are summarized in Table 3, which shows that, in addition to the expected significant between sessions main effect, due to practice, there are significant differences between the three groups in their average levels, thus confirming the graphical impression from Fig. 1. There is, however, a significant group × session interaction also which was consequently used to test the significance of the main effects mentioned. The large variance attributable to individual differences is of no especial interest here. Further partitioning of residual variance in the inset of the table shows that while there is a significant effect due to periods, there is no significant interaction between group and period, from which it may be deduced that the groups do not differ significantly in respect of the slope and curvature components identified earlier.

In the second analysis of variance of the present data, the two schizophrenic groups, massed versus spaced, were compared. The normal spaced group was not included in these comparisons since it so clearly falls outside the range of all other groups, as may be seen from Fig. 1. To facilitate this analysis it was necessary to equate the two groups in respect of the number of observations (10 sec. periods) in each session, and this was done by discarding the last six of the thirty observations for the massed group, leaving twenty-four in each. It is not anticipated that this procedure, necessary for achieving orthogonality in this analysis, had any important effect upon the outcome. The results are summarized in Table 4, which shows that there are significant differences between the groups' average levels per session, including a groups × sessions interaction. As may be seen from Fig. 1, the group given spaced practice is distinctly better. There is also a highly significant interaction between group, session and period, indicative of differences between slope and curvature. Examination of this interaction shows that these differences are confined to Session I, in which, at first, there is little difference between the groups. During the latter part of the session the spaced practice group improves more rapidly than the massed group and finishes with a distinctly higher score. But no further changes occur in the subsequent sessions II and III, which do not differ significantly between themselves. Thus, the superiority of level of the spaced group is maintained as is the absence of further differences in the other components, slope and curvature.

Table 4. *Summary of results of analysis of variance of the transformed scores: spaced practice*

Source	D.F.	M.S.	F	Significance
A. All sessions combined				
Between groups (G)	1	77·40	10·6	**
Between periods (P)	23	1·17	2·2	**
G × P	23	0·32	0·6	N.S.
Within groups (Individuals)†	960	0·53	—	—
Subtotal	1007	—	—	—
B. Distinguishing sessions				
Between sessions (S)	2	152·36	491·5	***
G × S	2	0·60	3·3	*
P × S	46	0·44	2·4	**
G × P × S	46	0·18	2·0	***
Residual‡	1960	0·09	—	—
Subtotal	2016	—	—	—
Total	3023	—	—	—

† Including individuals × periods interaction. ‡ including other interactions with individuals. N.S. = not significant. * = significant at the 5 % level. ** = significant at the 1 % level. *** = significant at the 0·1 % level.

DISCUSSION

The most important single finding emerging from this experiment is the absence of any significant reminiscence effect among psychotic groups given massed practice. This has been confirmed by others using schizophrenic subjects on the same task, administered and scored in comparable fashion (Claridge, 1960; Ley (in Eysenck, 1961); Rachman, 1963). It may be taken as established, therefore, that this absence of reminiscence is a real phenomenon and deserving of explanation, at least so far as schizophrenics are concerned. Within the reference terms of the inhibition concepts we have been dealing with, the most obvious explanation is that reactive inhibition failed to decay during the rest interval and so psychotics resumed practice at the previous level of performance. This explanation illustrates the crudity of our general initial hypothesis of a greater amount of reactive inhibition among schizophrenics. At least three factors governing reactive inhibition can ideally be distinguished. Schizophrenics may differ from normals in respect of the rate of growth of reactive inhibition, in the level ultimately reached, and in the rate of its spontaneous decay. The absence of reminiscence suggests that the greatest effect is related to the last of these three factors—a decreased rate of decay. This explanation can then be applied to the results obtained from the schizophrenics working under spaced practice, which resemble those for massed practice so much because the practice was insufficiently spaced for the reactive inhibition to dissipate during the rest periods allowed in the testing procedure. Consequently the development of conditioned inhibition in the manner described before was enhanced. This conditioned inhibition accounts for the depression in performance, and its extinction on resumption of spaced practice after each long rest interval, giving the warm-up or extinction increment effect observed in this group's results at the start of Sessions II and III.

In his work with psychotics, Venables has also employed the hypothesis that schizophrenics are subject to exaggerated reactive inhibition (Venables & Tizard, 1956), and has also identified its rate of dissipation as the important factor (Venables, 1959). He showed that the performance of chronic schizophrenics on a continuous reaction-time task fell off significantly faster than that of depressives. When faster responding was required, both depressives and normals showed the same rate of fall off as the chronic schizophrenics. Venables suggests that the inability of schizophrenics to dissipate reactive inhibition as fast as others put them at a selective disadvantage in the task where a slower rate of response was required.

The hypothesis encounters at least two difficulties in the present experiment, however. The first is the presence of the warm-up effect in the massed practice of psychotics in both Sessions II and III. If we postulate that reactive inhibition has not dissipated during the rest periods, how then can we use its absence to account for the extinction increment effect during the first few trials of the resumed practice? Possibly the crucial difference lies in the amount of reactive inhibition decayed: it may not be sufficient among schizophrenics to allow reminiscence to appear, but still sufficient to permit the extinction of conditioned inhibition. The somewhat shorter extinction increment discernible in the psychotics' massed practice curves (Fig. 1) is in keeping with this possibility. The second difficulty is the fall off, significant in the first case (Table 2), in the spaced practice of the schizophrenics during the 10 min. rest intervals—almost a 'negative reminiscence effect'. Perhaps the effect of some undissipated reactive inhibition is still persisting here also. Thus the explanation for our findings in terms of differential decay of reactive inhibition is not altogether satisfactory.

There are, however, equally serious objections to other possible explanatory hypotheses which, broadly speaking, are two in number. The first concerns differences in motivation, and the second in attention or set. Some aspects of the first can be thought of in terms of an additional variable—drive or 'D' in the Hullian equation. If we postulate merely a lower motivation among the psychotics than among the normals, this would permit the latter a higher build up of reactive inhibition (negative drive), before it counteracts the positive drive sufficiently to enforce an involuntary rest pause. Thus schizophrenics' performance levels might be lower, as was found, not only because the lower drive level would have the usual effect of failing to energize the habit strength developed to the same extent as among normals, but also because the more frequent involuntary rest pauses would further depress performance. But when the 10 min. rest pause came, reactive inhibition would dissipate, allowing the normals' reminiscence effect to be revealed, whereas the onset of the rest pause would find the psychotic with little or no reactive inhibition undissipated, and hence no reminiscence effect to be revealed by it. The extinction increment might still be present in their massed practice, however, because some conditioned inhibition would have developed in the normal way. This explanation, therefore, has the merit of avoiding the difficulty associated with the extinction increment in performance found at the beginning of Sessions II and III which the previous hypothesis discussed left unresolved. Turning to spaced practice, the differential motivation might reveal itself more fully than with massed practice, as neither normals nor schizophrenics could accumulate sufficient reactive inhibition to allow involuntary rest pauses to occur, and consequently the action of the higher drive would be the direct one of increasing the habit strength. That is, the action of motivation would be uncomplicated by reactive inhibition. Support for this view may be derived from Eysenck & Maxwell's identification (1961) of reminiscence as a measure of drive: the absence of reminiscence among schizophrenics would imply lowered motivation. But invoking drive differences does not account for either the 'negative reminiscence effect', nor the extinction increment observed with the spaced practice, so that here again the explanation fails to encompass the findings. While our data do not permit us to assess the relative merits of explanations based on the concepts of Hullian reactive inhibition and of drive, the theoretical formulations have the merit of allowing testable deductions to be made from them.

The third hypothesis considered concerns postulated differences in learning due to the nature of the patient's illness: these relate to differences in perception and in attention or set. Walton (personal communication) suggested that the perceptual organization of schizophrenics might be so disturbed as to fail to provide a sufficiently unchanging stimulus situation for reactive inhibition to develop, hence the absence of reminiscence. Alternatively, attention may be disturbed: patients may be preoccupied by their illness, be distracted by internal voices, etc., and be generally unable to concentrate. These might result in a failure to resume the task as quickly as normals do, so that the normal loss of set is exaggerated and the time taken for the true warm-up effect to occur is longer than the first few seconds of resumed practice to which it is normally confined (Eysenck, 1956*b*). This postulated loss of set would be particularly important in the case of spaced practice where the subject was continually having to stop and start again. The pursuit rotor may not be the best task for investigating this hypothesis and the present data do not provide any clear leads. The work of Rodnick & Shakow (1940) on reaction times supports this view, but without, however, excluding the other possibilities discussed. We have given these hypotheses less attention than they perhaps deserve, as they are not yet formulated in terms related in any rigorous way to learning theory concepts like inhibition and drive strength. It may well be that all the possibilities we have canvassed as relevant to the differences found between normals and psychotics—that is, differential decay of reactive inhibition, motivational differences, and differences in perception and set—may be operative.

We are indebted to Prof. H. J. Eysenck for his generous assistance, especially in making the apparatus and the data for normal subjects available to us. The co-operation of the Medical Superintendent of Springfield Hospital, Dr H. C. Beccle, and the medical and nursing staff is gratefully acknowledged. Some of the analyses reported were performed by Dr P. Slater, to whom we are doubly indebted for his permission to quote them.

REFERENCES

ADAMS, J. A. (1961). The second facet of forgetting: a review of warm-up decrement. *Psychol. Bull.* **58**, 274–98.

AMMONS, R. B. (1951). Effects of pre-practice activities on rotary pursuit performance. *J. Exp. Psychol.* **41**, 187–91.

BILODEAU, E. A. & BILODEAU, I. McD. (1961). Motor-skills learning. *Ann. Rev. Psychol.* **12**, 243–80.

BORING, E. G. (1913). Learning in dementia praecox. *Psychol. Monogr.* **15**, no. 2, 1–101.

BROADHURST, A. (1958*a*). Experimental studies of the mental speed of schizophrenics: I. Effects of a stimulant and a depressant drug. *J. Ment. Sci.* **104**, 1123–9.

BROADHURST, A. (1958*b*). Experimental studies of the mental speed of schizophrenics. II. Effects of practice. *J. Ment. Sci.* **104**, 1130–5.

BROADHURST, A. & BROADHURST, P. L. (1959). An analysis of the pursuit rotor learning of psychotics under massed and spaced practice. *Bull. Brit. Psychol. Soc.* **37**, 2*A* (Abstract).

CLARIDGE, G. (1960). The excitation-inhibition balance in neurotics. In *Experiments in Personality*, vol. II. *Psychodiagnostics and Psychodynamics* (ed. H. J. Eysenck). London: Routledge and Kegan Paul. Pp. 107–55.

EYSENCK, H. J. (1956*a*). Reminiscence, drive and personality theory. *J. Abnorm. Soc. Psychol.* **53**, 328–33.

EYSENCK, H. J. (1956*b*). 'Warm-up' in pursuit rotor learning as a function of the extinction of conditioned inhibition. *Acta Psychol.* **12**, 349–70.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria: An Experimental Application of Modern Learning Theory to Psychiatry.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1961). Psychosis, drive and inhibition: a theoretical and experimental account. *Amer. J. Psychiat.* **118**, 198–204.

EYSENCK, H. J. & MAXWELL, A. E. (1961). Reminiscence as a function of drive. *Brit. J. Psychol.* **52**, 43–52.

FELDMAN, M. P. (1963). A reconsideration of the extinction hypothesis of warm-up in motor behavior. *Psychol. Bull.* **60**, 452–9.

GRIEW, S. & LYNN, R. (1960). Construct 'reactive inhibition' in the interpretation of age changes in performance. *Nature, Lond.*, **186**, 182.

HULL, C. L. (1943). *Principles of Behavior.* New York: Appleton-Century.

HUSTON, P. E. & SHAKOW, D. (1948). Learning in schizophrenia: I. Pursuit learning. *J. Personality*, **17**, 52–74.

IRION, A. L. (1949). Reminiscence in pursuit-rotor learning as a function of length of rest and amount of pre-rest practice. *J. Exp. Psychol.* **39**, 492–9.

JENSEN, A. R. (1961). On the reformulation of inhibition in Hull's system. *Psychol. Bull.* **58**, 274–98.

KRAEPELIN, E. (1913). *Psychiatrie* (Vol. 3). Leipzig: Barth.

LYNN, R. (1962). Aging and expressive movements: an interpretation of aging in terms of Eysenck's construct of psychoticism. *J. Genet. Psychol.* **100**, 77–84.

Rachman, S. (1963). Inhibition and disinhibition in schizophrenics. *Arch. Psychiat.* **8**, 91–114.

Rodnick, E. H. & Shakow, D. (1940). Set in the schizophrenic as measured by a composite reaction time index. *Amer. J. Psychiat.* **97**, 214–25.

Sandler, J. (1955). A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*. *Brit. J. Psychol.* **46**, 225–6.

Venables, P. H. (1959). Factors in the motor behavior of functional psychotics. *J. Abnorm. Soc. Psychol.* **58**, 153–6.

Venables, P. H. & Tizard, J. (1956). Performance of functional psychotics on a repetitive task. *J. Abnorm. Soc. Psychol.* **53**, 23–26.

(*Manuscript received* 5 *July* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 333–354
Printed in Great Britain

A STUDY OF 200 SCHOOL CHILDREN BY MEANS OF THE TEST OF FAMILY ATTITUDES

By LYDIA JACKSON

370 *Russell Court, Woburn Place, London, W.C.* 1

The original study which served as a basis for 'a test of family attitudes' was limited to 110 children between the ages of 6 and 12: 40 normal, 40 neurotic and 30 delinquent. In the present study an attempt is made to give the 'test' a broader basis and to establish norms by testing 200 children from ordinary schools (100 boys and 100 girls between the ages of 8 and 12), who were not known to be disturbed. The responses were analysed in somewhat greater detail than in the previous study. The results showed no discrepancy with the original study and may be considered as confirming it. Tables are offered for use in comparing the responses of children who may be emotionally disturbed and whose family relationships need to be investigated but can be approached only indirectly. Four complete sets of stories, two by normal and two by disturbed children, are attached as illustrations.

INTRODUCTION

This study was undertaken with a view to giving a broader basis to the original piece of work which showed some significant differences in the responses made by normal and by disturbed children to a series of specially designed pictures. It was hoped to establish norms with which the responses of the disturbed children could be compared.

Two hundred children between the ages of 8 and 12 were tested by the author and a colleague, an educational psychologist, with the help of several students of psychology. The children came from primary and secondary schools and were taken for testing at random. There were 50 boys and 50 girls between the ages of 8 and 10, and 50 boys and 50 girls between the ages of 10 and 12. No intelligence tests were given by the investigators, and it was assumed that the great majority of the children were in the average range, as the headmasters were always asked to supply 'normal children' or 'children without problems'.

The expectation, borne out by the previous study (Jackson, 1950), was that the normal subjects would give a preponderant number of 'positive' responses of a realistic type, which could be interpreted as reflecting satisfactory relationships between the child tested and his family and friends; that sado-masochistic responses would be few, and that accidents, crime, death and gruesome fantasy would come into the stories only occasionally. With a view to classifying the subjects' responses, the stories produced were broken up into a number of items. Thus, picture 0 stories were broken up into 17 items, picture 1 into 25, picture 2 and 2A into 24, picture 3 into 18, picture 4 into 16, picture 5 into 17 and picture 6 into 17. These items were designed to cover the contents of the children's compositions as fully as possible, while at the same time allowing for them to be treated statistically. Subsequently, each of the 1400 stories was analysed in order to see how the content of each one corresponded with the items obtained from the material of all the stories about the same picture. The analysis was made by the writer, with the help of a colleague who had also assisted in giving the Test, and analysed 25 sets of stories taken at random. It goes without saying that the assessment of some items, such as 'punishment severe' or 'naughtiness minor', were bound to involve subjective judgement, and some disagreement between the two judges did arise in a few cases on such items. But on the great majority of items there was agreement between the two. The results of the analysis are shown in Tables 1–7.

It is evident from these tables that the majority of the children tested saw the adults and children in these pictures as belonging to a family unit and that they saw family relationships in a realistic, matter-of-fact way. This can be seen in their judgement that, although the child in the story is described as 'good' or 'nice', he cannot be good all the time, so that occasional 'naughtiness' or 'badness' is accepted as a fact of life and is usually regarded as deserving some punishment. A very common description of the parent–child relationship is that parents like the child when he is 'good' and cease to like him when he is naughty, while the child's positive feeling for parents is suspended during and immediately after punishment. More than half the children tested had little hesitation in recognizing jealousy, envy and feelings of insecurity in the older child confronted with a new baby, and expressed the belief that parents preferred the baby to the older child. The remainder, however, believed that the parents liked both children equally. Punishment for 'naughtiness' was usually moderate, such as 'smacking' or 'sending to bed', very rarely 'thrashing' or depriving of meals for the whole day.

When dealing with pictures 3, 4 and 5, where 'naughtiness' and violence are deliberately suggested, most of the children tested were able to take this in their stride, without undue anxiety or eagerness. Half of them said that the child in picture 3 was alone because he had been 'naughty', but well over half gave the story a satisfactory ending, such as forgiveness (by the parents), reconciliation with them, or the reformation of the child.

In picture 4 the majority saw the situation as one of threat or danger, yet again more than half concluded the story with a 'happy' ending. In picture 5 more than half saw the situation as picnic or play and the child as being helpful or innocuously occupied, and a considerable majority said that the mother (who is not in the picture) was available to the child. In picture 6 half the subjects said that the parents were discussing the child, but the majority did not think the child had been 'naughty', and about two-thirds gave the story a positive ending.

An interesting feature of the tables is that the variations in the number of identical responses to the same item by different age-groups and different sexes are very small; in many cases the numbers are practically identical.

A detailed analysis of story content follows.

ANALYSIS OF STORY MATERIAL

Picture 0. A large female figure bending over a cradle (designed to suggest maternal protection and the child's dependence on the mother).

The stories about this picture rarely have a plot or are really imaginative.

If there is *a plot*, it is frequently taken from the fairy tale about a mother who badly wanted a baby and whose wish came true either in the normal way, or through finding a baby on her door-step.

Table 1. *Analysis of responses. Picture* 0

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	Mother to the baby (or near relative; grandmother, aunt)	43	45	36	44	168
(2)	Not mother (baby sitter or other stranger)	7	2	14	6	29
(3)	Baby 'nice' or 'good'	34	36	30	35	135
(4)	Baby naughty or 'bad'	6	3	11	6	26
(5)	Baby always nice	7	9	5	3	24
(6)	Baby always naughty	0	0	0	0	0
(7)	Baby in-between (sometimes naughty)	37	32	33	35	137
(8)	Mother's reactions: comforting, helpful	18	12	23	19	72
(9)	Mother's reactions: punishing or restraining	20	21	18	19	78
(10)	Mother's punishment light ('tells off', puts to bed)	17	20	18	17	72
(11)	Mother's punishment severe (smacks hard, deprives of food)	2	0	1	0	3
(12)	Mother likes the baby a lot	40	43	45	45	173
(13)	Mother does not like it very much	6	2	0	2	10
(14)	Father benevolent	32	31	31	28	122
(15)	Does not like the baby	2	0	1	4	7
(16)	Father's feelings towards baby not mentioned	15	12	16	15	58
(17)	No father	2	6	1	4	13

Table 2. *Analysis of responses. Picture* 1

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	Child 'good' or nice	23	25	29	23	100
(2)	Child naughty or 'bad'	10	14	9	16	49
(3)	Child good always	10	9	3	1	23
(4)	Child naughty always	1	3	1	2	7
(5)	Child in-between (sometimes naughty)	36	35	38	41	150
(6)	Child's 'naughtiness' serious (destroys wilfully, steals, fights parents)	0	0	1	1	2
(7)	Child's 'naughtiness' 'normal' (disobeys, plays tricks, breaks things by accident)	46	36	40	47	169
(8)	Parents punish severely (thrash, deprive of food for more than one meal, or of treats for long periods)	1	2	0	3	6
(9)	Parents punish lightly (smack, send to bed, 'tell off')	36	36	39	37	148
(10)	Parents like the child very much	39	40	41	45	165
(11)	Parents like the child not very much	8	8	5	5	26
(12)	Parents like the child always	12	16	5	12	45
(13)	Parents like the child not always	23	22	22	24	91
(14)	Child likes the parents very much	35	40	38	45	158
(15)	Child likes the parents not very much	6	5	4	5	20
(16)	Child likes the parents always	8	12	10	16	46
(17)	Child likes the parents not always	12	23	20	20	75
(18)	Man outside is a relative or friend	38	42	40	33	153
(19)	Man outside is benevolent	35	39	31	31	136
(20)	Man outside is a stranger	12	8	7	16	43
(21)	Man outside is hostile or dangerous	8	5	4	7	24
(22)	Man outside likes the child	35	36	31	37	139
(23)	Man outside does not like the child	11	9	8	6	34
(24)	The child likes the man	32	33	36	30	131
(25)	The child does not like the man	10	11	4	7	32

The baby's *father* is rarely mentioned spontaneously. When question (ix) ('Tell me something about the baby's father') is put, it is usual for the child-subject to describe his father's appearance or to mention his occupation, or both. The *father's feelings* towards the baby are rarely mentioned spontaneously, but occasionally the subject says that the father 'made the cot for the baby', or 'plays with the baby'.

With a very few exceptions *the mother* is said to want the baby and to be very fond of it.

The *baby's naughtiness* is very often equated with 'normal' features of a baby's behaviour, such as 'crying' (most often); 'kicking off clothes and bed covers', 'throwing things out of the cot', 'tearing things up'. Sometimes it is made explicit that crying is 'naughty' because it keeps the parents awake.

Occasionally the subject seems to forget that the baby is very young and ascribes to it 'naughty' activities more appropriate to an older child, such as 'looks in drawers and pulls things out', 'plays up, wanders round the house and tips up cushions and things like that', or even 'climbs trees'. This points to identifying the baby with himself, or with a 'naughty' younger sibling who is no longer a baby.

Mother's reactions to baby's 'naughtiness': in the great majority of cases the mother is said to '*smack*'—'not very hard' is sometimes added. Less frequently the mother '*tells off*'. Very occasionally the subject remarks that the baby is 'too young to be smacked', but apparently not too young to understand 'telling off'. Occasionally mother merely restrains: 'puts him back into the cot', and in a few stories she reacts positively, by 'picking the baby up', 'cuddling him' or 'giving him milk to drink' or a 'dummy'.

In a few cases *the story is extended* to include the baby growing up, perhaps even marrying and helping the parents.

Picture 1. A man and woman sitting on chairs side by side, turning their backs on a child who is sitting on the floor. Outside a closed door a male figure is standing (designed to suggest the child's exclusion from intimacy between the parents with the consequent threat to its security)

A simple plot occurs more frequently in stories about this picture than in stories about picture 0.

The couple sitting on chairs are almost invariably described as 'watching television'. Very occasionally they are said to be 'waiting in a doctor's surgery'.

The child on the floor. Usually described as 'playing with his toys'. Very occasionally is seen as having a *damaged* (*broken*) *leg*. Sometimes described as 'the baby' (probably a carry-over from the previous picture). When age is given it is usually under 8. Nearly always his behaviour is described as 'half-and-half', 'sometimes naughty, sometimes good'.

Identification. There is no consistency in relation to sex. A boy may see the child as a girl or a boy. The same is true of girls. There may be a link with their having sibs of one or the other sex, but this has not been explored; and it is doubtful that it would prove significant.

Naughtiness. Examples given: breaks windows (on purpose); 'messes about'; uses milk bottles as targets for catapults; teases the dog; plays tricks on the family; gets into a temper and tears things up; gets muddy playing games; makes marks on the wall (inside the house) with dirty hands; sulks; jumps on furniture; looks through drawers in mother's bedroom; pulls up flowers in the garden; splashes water in the kitchen; throws pillows at elder brother; climbs over chairs; argues with mother; throws stones at other boys; fights; plays on beds; shoots water with water pistol; slams the door or leaves it open; switches television off and on; slides down banisters; takes biscuits or jam without permission; pulls pillows out of pillow cases; is rude to mother and father.

Attitude to parents. Is usually said to like or 'love' the parents, but *not* when they punish him or 'tell him off'. In a few cases he is said to love his parents *always*. Very occasionally the child is said to prefer one parent (usually the mother).

Attitude to 'uncle' or 'elder brother' (the man outside, see below).

Usually 'cupboard love'. The child is said to love the uncle 'because he brings him presents' or 'because he gets him out of trouble', or 'takes him out to the park and gives him rides in the car'.

Table 3. *Analysis of responses. Picture* 2 *or* 2A

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	Older child 'good' or nice	24	29	23	30	106
(2)	Older child naughty or 'bad'	11	6	12	9	38
(3)	Older child always good	6	14	9	9	38
(4)	Older child always naughty	2	1	3	1	7
(5)	Older child in-between (sometimes naughty)	35	27	30	31	123
(6)	Older child's naughtiness serious (injuring baby, destroying etc.)	2	1	2	4	9
(7)	Older child's naughtiness minor	38	26	31	27	122
(8)	Older child likes the baby	11	19	16	15	61
(9)	Older child does not like the baby	8	11	12	9	40
(10)	Older child's thoughts about the baby hostile	4	9	11	8	32
(11)	Older child's thoughts reflecting insecurity at loss of attention	10	12	21	24	67
(12)	Older child's thoughts, wish to be the baby	1	5	3	1	10
(13)	Older child's thoughts, jealousy expressed as such	9	12	16	17	54
(14)	Older child's other thoughts or reactions positive (playing with baby, helping mother)	24	20	11	18	73
(15)	Older child's reactions negative (trying to harm baby or parents, damage home, etc.)	3	2	8	8	21
(16)	Older child running away	3	5	2	7	17
(17)	Parents' reactions; positive (reassurance, help)	3	6	11	18	38
(18)	Parents' reactions; negative (punishing, sending away)	33	18	30	16	97
(19)	Baby same sex as older child	22	19	30	15	86
(20)	Baby opposite sex to older child	22	30	15	31	98
(21)	Parents love the older child more than the baby	1	7	1	2	11
(22)	Parents love the baby more than the older child	27	29	26	24	106
(23)	Parents love them both the same, or child aware that present marked attention to baby is necessary	23	18	28	26	95
(24)	All ends well	9	17	25	18	69

If the man is seen as 'burglar', the child sometimes warns the parents and gets rewarded. If an 'elder brother', joint activities are sometimes mentioned. The theme of a hide-and-seek game occurred more than once.

The parents

Most frequently like the child 'when he is good'. In a few cases it is made explicit that the father is stricter than the mother. The majority of the children state that the parents 'do not like the child when he's naughty'. Very occasionally the parents are said to 'love the child always, whether she's naughty or good—because she's their daughter' (example from an only child).

Punishments

The following examples are given: 'smack him'; 'tell him off'; 'send or put him to bed'; scold him; 'send him to bed without any tea'; 'forbid him to watch television for a week'; 'do not let him go out to play'; 'give him a good hard smack' (rare): 'a little smack'; 'belt him' (only one example).

The man outside

Nearly always seen as 'uncle' coming to visit; as a rule, benevolent, bringing presents; sometimes neutral (neighbour coming for a cup of tea, a father's friend); sometimes a postman or 'the rent man'; rarely dangerous or malevolent, e.g. a burglar or kidnapper. On a few occasions seen as 'the doctor'. More than once said to strike the child with the door and hurt him, so that he had to be taken to hospital.

Table 4. *Analysis of responses. Picture* 3

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	Child alone because it has been naughty	23	21	32	31	107
(2)	Child alone because it is bereaved, or parents ill	6	9	5	7	27
(3)	Child alone because it is left alone in the house by parents	9	12	7	9	37
(4)	Child alone because it chose to be alone	17	11	8	3	39
(5)	If naughty, is repentant	8	7	11	15	41
(6)	If naughty, wishes to do it again	4	1	3	4	12
(7)	If naughty, plans to escape	7	12	8	2	29
(8)	If naughty, is longing for parents	6	7	7	8	28
(9)	If naughty, is day dreaming	25	16	19	13	73
(10)	Parents gone out	12	15	12	7	46
(11)	Parents are in the house (next room, downstairs)	30	21	31	22	104
(12)	Door of the room locked	18	15	20	12	65
(13)	Door of the room just shut	25	33	23	26	107
(14)	Child left alone a long time (a day, many hours)	9	11	15	8	43
(15)	Child left alone a relatively short time	38	34	30	33	135
(16)	End of the story positive (reconciliation, forgiveness, reformation of the child, restitution, recovery)	31	27	37	34	129
(17)	End of the story negative (further deprivation, punishment, rejection)	5	2	3	4	14
(18)	End of the story inconclusive	13	21	6	11	51

Table 5. *Analysis of responses. Picture* 4

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	The situation seen as play (or anything non-threatening)	19	22	20	21	82
(2)	The situation seen as threatening	21	28	30	28	107
(3)	Male figure seen as father of the child	43	43	40	40	166
(4)	Male figure seen as stranger	9	9	9	10	37
(5)	Male figure seen as benevolent	20	19	17	20	76
(6)	Male figure seen as threatening or dangerous	27	29	25	24	105
(7)	Female figure seen as mother of the child	46	40	42	45	173
(8)	Female figure seen as sister or other relative	2	5	2	4	13
(9)	Female figure seen as stranger	2	4	3	1	10
(10)	Female figure seen as protecting the child	18	13	29	20	80
(11)	Female figure seen as offering it as victim for punishment	7	9	4	5	25
(12)	Adult figures both approve of what is happening	23	24	16	24	87
(13)	Woman and man opposed to one another	25	25	32	26	108
(14)	The story has a positive ending (child forgiven, rescued or successfully protected)	27	19	28	32	106
(15)	The story has a negative ending (child rejected, maltreated or killed)	5	2	2	1	10
(16)	The story ends inconclusively	18	29	17	17	81

Table 6. *Analysis of responses. Picture* 5

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	The situation seen as picnic or play	25	33	19	31	108
(2)	The situation seen as dangerous or threatening	24	17	29	17	87
(3)	The child is being 'naughty' (doing something forbidden)	18	15	31	17	81
(4)	The child is being helpful, playing or occupied innocuously	30	35	17	27	109
(5)	Male figure seen as father of the child	35	35	30	35	135
(6)	Male figure seen as elder brother	2	3	5	4	14
(7)	Male figure seen as stranger	13	9	15	12	49
(8)	Male figure seen as benevolent	28	28	21	27	104
(9)	Male figure seen as threatening or dangerous	19	18	23	19	79
(10)	Child's mother available	42	43	45	38	168
(11)	Child's mother not available or dead	7	7	5	10	29
(12)	Child's mother if available, comforting and approving	18	20	19	19	76
(13)	Child's mother if available, disapproving	13	8	13	7	41
(14)	Child's mother if available, punishing	4	7	6	2	19
(15)	The ending of the story positive (reward of the child, if good, reconciliation, reformation, forgiveness)	25	22	31	25	103
(16)	The ending of the story negative (punishment, rejection, sending away)	4	3	3	2	12
(17)	The ending of the story inconclusive	20	25	12	17	74

Table 7. *Analysis of responses. Picture* 6

Item no.	Responses	Boys 8–10	Girls 8–10	Boys 10–12	Girls 10–12	Total
(1)	Parents quarrelling	9	9	10	10	38
(2)	Parents discussing the child	25	27	26	26	104
(3)	Parents trying to solve a problem not directly related to the child	17	20	18	18	73
(4)	Child has been naughty	13	16	15	18	62
(5)	Child has not been naughty	37	34	32	31	134
(6)	Child does not understand what parents are discussing	19	16	17	11	63
(7)	Child understands	24	28	29	34	115
(8)	Child is anxious	15	10	12	13	50
(9)	Child criticizes parents	4	3	1	9	17
(10)	Child tries to be helpful	8	13	5	8	34
(11)	Parents are loving to the child	17	24	14	20	75
(12)	Parents are hostile or rejecting	3	6	5	4	18
(13)	Parents find a solution to the problem	21	27	13	18	79
(14)	The ending of the story positive (things sort themselves out; child is forgiven or reforms; problem is solved)	38	38	39	40	155
(15)	The ending of the story is negative (problem remains unsolved, child is rejected or sent away)	0	5	2	1	8
(16)	The end of the story is inconclusive	12	7	9	9	37
(17)	One parent loving, the other hostile	7	1	6	8	22

Picture 2 *and* 2A. A man and a woman with a baby in her arms sitting side by side on a sofa. A small boy (girl in 2A) standing some way off, gazing at them (designed to suggest the problem of adaptation to a sibling, with all that it may involve)

The older child. Character. Usually 'half-and-half'; often 'rather naughty'. Sometimes said to be 'spoilt'; occasionally described as 'adopted' or as the 'neighbour's little boy/girl'. Change in character from 'good' to 'naughty' is sometimes explicitly linked up with the baby's arrival and the feeling of being no longer wanted, or neglected.

Naughtiness. Disobeys parents; sulks; won't eat his meals; makes a lot of noise; hits other boys and kicks them; steals things from shops; digs up plants in the garden; climbs over furniture in dirty shoes; makes a draught by leaving doors open; plants stink bombs under kitchen stove; cuts chairs; cuts dad's slippers with scissors; climbs walls and wears out the toes on his shoes; takes parents' cigarettes and smokes them (all from boys); hits people; hurts baby; knocks over things on purpose; won't run errands for mother; won't go to bed at night; won't have her bath; hides the baby from parents; gets in her mother's way and breaks things; is rude to people; plays in the street when not allowed to; is defiant (these from girls).

Reactions to baby

(*a*) *Positive.* Thinks it a nice baby; wants to hold it; looks after it well when parents out; accepts the baby being liked more than himself/herself; is concerned at baby's crying or being hurt; looks forward to baby growing up and being a play companion; plays with baby; pleased at baby's sex (boys seem to prefer brothers); wonders what it is going to be called; thinks the baby is lovely; looks forward to taking the baby out; to having a baby of her own (girls); thinks it is her *own* little brother.

(*b*) *Negative.* Teases baby; hides baby from parents; prevents it going to sleep by making a noise; is 'very jealous' of it; 'despises it'; would like to 'wring its neck'; shouts at baby; makes it cry; hits it; knocks it about (boys); resents limitations imposed for its sake, such as keeping quiet; runs away from home because cannot bear baby being preferred; cries out 'nobody wants me any more' (girls); wonders 'why should the baby be cuddled and not me?'; wonders 'why couldn't it happen to me?' (i.e. having a baby) (girls); pinches the baby; keeps on hitting it; breaks things and pretends the baby's done it; takes baby out and leaves it outside a shop; tips the pram so that the baby can fall out; wishes the baby 'had never been born'; thinks 'what a spoilt baby!'; does 'something wrong' and blames it on the baby; wakes the baby up; 'makes the baby bad' (like himself); hopes the mother will give the baby a smack; shakes the cot; breaks the baby's doll; thinks 'I wish I could hit the baby'; thinks it is 'going to be horrible' when it grows up.

Reactions to parents

(*a*) *Positive.* Takes them to live with him when they grow old (despite running away from them because they loved the baby more); finds (on running away) that he 'couldn't do without his mother and father'; wants to sit in the armchair between father and mother; lays table for dinner for the family; helps the mother in the house; does shopping for mother.

(*b*) *Negative.* Thinks they like the baby best (most often); pretends he/she is hurt to draw parents' attention to himself (herself); thinks 'it isn't fair' that parents love the baby more; does not know they love them equally; 'wonders' whether the parents would love him/her as much on account of having a new baby; feels not wanted any longer; not loved any more; wants 'a lovely lollipop'; wants a new bike; wants money for sweets; feels 'left out'; wants an electric train set; calls parents names ('bighead' to father); makes mother angry (deliberately) by waking the baby; does not like her mother because she would not let her hold the baby; does not love her mother any more.

Parents' attitude to older child

(*a*) *Positive.* Ask 'why the child is weeping, or sulky, or why it had run away' and compensate; pet her or love her; say they are sorry and promise to love her the same as baby; say they like her; explain why the baby had to be looked after more; think her helpful and that she will be a good housewife when she grows up; like her better than the baby who pulls her hair and is 'cruel'; 'like her best because they had her longest'.

(*b*) *Negative*. When the older child is naughty, parents smack him/her, tell him off; send him to bed without tea or supper; do not 'trust him with the baby'; do not give him pocket money; put him into his room; father canes him (one example); hit her (girl) when she hits the baby; make example of the (good) baby; tell him to leave the baby alone; send him out of the room; have no time for the boy; send away to a Home for two years; father refuses to go out with the boy because of the baby; father tells the girl 'to be quiet', or to 'wash herself'; take no notice of her.

Parents' attitude to baby

Admire it; make a fuss of it because it is small, soft, pretty, 'gooder' than the older child; like the baby best (because it is not so naughty, because it doesn't make so much noise; because it is 'kind', because it does not 'get up to so many tricks'; because it 'doesn't roll in the mud'); they are 'delighted with it'; they wonder what to call it. Very occasionally worried because the baby is not well. Baby 'never breaks things or pulls them (plants) up'.

Conclusion

(*a*) Child: 'grows out' of his/her hatred, jealousy or dislike of the baby and begins to love it; they become good friends and play together. If sent away, comes back reformed, no longer naughty. 'Gets over' jealousy when she is older and 'is kind' to her brother.

(*b*) Parents: realize their mistake (in showing too much love to the baby) and compensate the older child, or correct their ways and 'love him/her as much as the baby'. Occasionally, parents are said to share the care of the two children: mother looks after the baby, father after the older child. When the baby is older, they would like both children the same.

Picture 3. A child sitting alone on a stool, with a door closed (designed to suggest transgression and the ensuing loneliness)

There are four main variants of story about this picture. These are in order of frequency: (*a*) the child is at home and has been naughty; (*b*) is at home alone because the parents had gone out; (*c*) is at school, put outside the classroom for naughtiness; (*d*) is an orphan in an orphanage or in a house by himself. Occasionally he is in a doctor's surgery; once or twice only he is sitting outside a sick room, waiting for doctor's report on sister or mother.

Reasons for being alone

(*a*) *Not naughty*: practising singing; has nothing to do; wonders what he can do; learning to count; no one to play with; watching television; just thinking; likes it; parents out; waiting for mother's return from shopping; waiting for father's return from work to get his pocket money; because baby is being washed in the next room; in surgery waiting for an injection.

(*b*) *Bereaved*: no brothers or sisters; no parents; has no friends; is an orphan in an orphanage; grieving for a pet animal that had died; parents dead; outside the bedroom of a sick mother or sister, or baby; is kidnapped; abandoned by parents.

(*c*) *Naughty*: unspecified; thinking up mischief; punished by teacher at school; sulking because he had been made to eat cauliflower; because frustrated by parents refusing toys and chooses to withdraw from them and even lock himself up; planning to run away; because he doesn't like anyone.

Length of time during which the child remains alone

Anything from 5 minutes to a month (when kidnapped), but the most usual length of punishment by isolation is from half an hour to an hour; 'about three hours' when parents are out at a dance.

Type of naughtiness

Spilled paint on new carpet; turned off gas in the kitchen; stole money from mother's purse and spent it on sweets; eaten up mother's cakes and tarts; played out late; did not tidy up his room; brought mud into the house and put it all over the sofas; painted when forbidden to do so; broke legs off chairs; won't drink his milk; sucks his thumb; slapped mother's face and kicked father; scratched the paint work; keeps on telling lies to his mother; hit the baby; did not feed his pet bird and frightened it so that it died; kicked father's beer over; was rude to guests; kept on hitting boys younger than himself; kept on slapping baby; stole 6*d*. from the purse of a visiting aunt; refused to go to bed when told; made faces at the teacher; threw ink-wells about and just missed the teacher's head; threw ink pellets about and hit the teacher; kept on talking (or shouting) in class; brought chewing gum to school and ate it in class; won't share a book with another child in class; drew faces on blackboard; tore all his clothes climbing; tore out toes of his shoes by climbing walls.

Child's attitude to his transgressions

Plans to escape; day-dreams of food and games; wishes he had not done it; wants to 'make up for it' (replace mother's broken vase); would like to hit (baby) again; thinks teacher is very mean; is very ashamed of himself; makes up his mind to be good; decides to be good and kind to parents; is feeling lonely.

Punishment meted out to child

A hard slap; sent to bed; locked up; is not given the toy his father bought him; no meals except water, or 'bread and butter'; no tea or supper (yet often is said to be given food later by mother); forbidden to go out to play; abandoned altogether by parents (one case).

Parents' positive attitude to child

Mother does not tell the child about visiting a sick friend 'not to upset him'; after child runs away and is brought back, mother 'kisses him and says "don't run away again"'; parents forgive him when he says he is sorry.

Conclusion

Parents forgive the child; child reforms, is 'never naughty again', 'never does it again'; is adopted by kind people (if bereaved); is re-united with the parents if separated.

Picture 4. A rather small woman holding up a child and a large man with his arm up facing them (suggesting the possibility of aggression and 'betrayal' by the parents)

In a great majority of cases the man is seen as the father of the child, but in a few cases as: kidnapper, an escaped convict, a fairy-tale giant, a 'funny man', a 'circus man', a 'park attendant', a 'forester', 'a layabout and thief', a 'fair-stall keeper'.

If *father*, he is very often 'a famous boxer' who is teaching his child how to box, or 'showing it his muscles'. Sometimes he is just 'playing with the baby'. In one or two cases the situation is described as 'having a picnic in the woods', or 'a game on the sea-shore'. In one case the man is said to be 'a doctor' (father) who is looking at the baby's rash.

Almost as frequent is the theme of *rejection* of the child by the father, who is said not to like babies, being too poor to keep another baby, or being disappointed in the baby's sex. The mother is usually described as pleading with the father on baby's behalf; sometimes she leaves the father because of the child, or conceals the child until it grows up and renders a service to the father who then accepts it.

In a relatively few cases the mother is described as cruel and wanting the father to hurt the baby; in one case she is said to be using the baby as shield to protect herself. In one case the scene is said to be happening 'on the stage', in another the child is said to be 'a dummy' on which the man practises boxing. In one case he is the father claiming his baby from a woman kidnapper. In one or two stories the mother is simply wanting the father to hold the baby, or asking him to wash it. More often she produces it (usually reluctantly) for father to reprove or to smack, for having done something 'wrong'.

In a few cases the female figure is described as a 'little girl' and the baby as 'a doll'.

Child 'naughty'

Been playing with father's (dangerous) tools; woke the father up by crying; wet his bed; 'done something wrong' (unspecified).

Child not naughty

Is being taught to box; to defend himself; is being shown father's muscles; is trying to hit back in play; is laughing at father.

Mother's attitude

(*a*) *Protective.* Stops play if the father gets 'rough'; holds the child up to show how small it is and that it must not be punched; to show how thin he is (because starving); threatens to leave father if he does not provide food and clothes; tries to mollify father; encourages child to learn how to fight; so that he can reach father's height; so that he can see what father is doing; likes him provoking father because she does not like father.

(*b*) *Punishing.* Holds the child up to be punished by father; co-operates with father in hurting the child because she rejects or hates the child; uses the child to protect herself from father's blows; hands the child over to a kidnapper, or a giant.

Father's attitude

(*a*) *Benevolent.* Showing muscles; having a game of boxing; just 'having a little game with the baby'; wants to hold the baby; wants the boy to be a boxer like himself; buys new clothes for the baby.

(*b*) *Hostile.* Does not like babies; does not want a girl (or a boy); wants to get rid of the child; is jealous of baby; does not like the way she is dressed; threatens to beat the mother up.

Conclusion

Parents (if poor) get wealthy and happy and keep the child; father forgives the child's naughtiness and does not punish it; warns it for the future; if both parents had been cruel, child gets adopted and refuses to return to them; father, after rejecting the child, accepts and likes it later. Child recovers lost parents. Man recovers (from madness), buys a house and the family live happily ever afterwards. Mother runs away from cruel husband, re-marries and the child has a kind step-father. Man is put in prison for cruelty.

Picture 5. A child crawling and stretching out towards a vaguely defined object, with a male figure standing behind with his arm raised (suggesting the attraction of forbidden fruit, and the likelihood of punishment)

In a great majority of cases the man is again seen as the father of the child, but occasionally as an elder brother, a boy companion, or even a stranger, sometimes hostile, sometimes helpful or benevolent.

There is a considerable variety of themes and interpretations of what the child is doing, but the themes of 'playing with the fire' or 'catching butterflies' are the most frequent ones. The male figure is usually restraining or protecting, sometimes threatening.

Unusual themes occurring only once in the total number of the stories: (*a*) girl escaping from a cruel step-father; (*b*) girl drifting in a boat, sending smoke signals; (*c*) girl burns father's examination papers unintentionally; (*d*) girl burns a compromising document relating to father in another man's house (by a child of divorced parents); (*e*) girl finds a cache of weapons in a gangster father's drawer, and he gets condemned and hanged; (*f*) father sets fire to grass by dropping a cigarette and the girl helps to put it out; (*g*) girl plays a game with father pretending to be animals; (*h*) child leaves both parents and lives in the hills alone; (*i*) child discovers oil in the ground.

Child 'naughty'

Making a fire in the garden; playing with fire in the home; is 'very stupid'; lights some leaves in the wood and sets it aflame; throws doll into the fire at home; sets garden on fire, then turns on water and floods it nearly drowning herself; refuses to come in when called; throws father's cigarette papers into the fire; sets her dress on fire; disobeys prohibition not to go out; plays with matches; fights with the 'boy'.

Child not 'naughty'

Is chasing butterflies; feeding birds; looking for firewood, stumbles and falls; trying to catch smoke from the fire; rescuing baby from the fire; saving father's wallet with £100 in it from the fire; watching birds; watching animals and birds she loves in a wood; discovers oil in the ground belonging to the County Council (and is rewarded); tries to retrieve a butterfly or bird from being burnt; is on a picnic with parents in the woods; looks for a bird that had flown away; is with father near a camp fire; is being rescued by a man from a drifting boat; is playing with father; would not box with the father for fun.

The man's role in the story

(*a*) *Protective or benevolent.* The man (seen *as father* of the child in the great majority of cases) prevents the child from climbing a tree ('might break her neck'); stops her from playing with the fire; rescues her from a drifting boat; helps her to catch butterflies (or birds); puts the fire out and rescues the child from being burnt; teaches the child to box, or tries to interest her in boxing.

(*b*) *Punishing.* Smacks hard with stick for lighting fire or playing with it; threatens to strike if she does not release the bird she had caught; stops pocket money in punishment; takes the disobedient child to the police and leaves her there; pushes the child into the fire; smacks and scolds.

(*c*) *Seen as 'brother' or 'boy'.* A friendly sibling playing with girl; a boy persecuting the girl who fights and knocks her down; as one of the two brothers who helps the brother with a broken leg; as a bully who takes money from her; as an avenging elder brother whom she had got into trouble (these variations are usually a single example of a theme).

(*d*) Seen as a '*giant*' who threatens to eat child (a single instance).

(*e*) Seen as a '*kidnapper*', wanting to take the girl away (a single instance).

(*f*) Seen as a '*gardener*' or '*park keeper*' forbidding the lighting of fires (two or three instances).

Mother's attitude

(*a*) *Benevolent or comforting.* She is fairly often represented as blaming father for being careless when he is left with the child and holds him responsible for accidents; is on the child's side when father punishes and 'tells him off' for smacking child; bandages the child's burned hand, or her pricks and scratches; if the male is a 'boy' bullying girl, mother reports him to his parents, or punishes him herself.

(*b*) *Disapproving or punishing*: withholds money (£2) to pay the fire brigade, called because child started a fire; tells her off, or puts her to bed for playing with the fire; blames the girl for letting a bird go; scolds her for hurting or catching butterflies or birds; whips boy and girl (brother and sister) for fighting; forbids the girl to go to the park again as she would not come home when called; supports father in punishing child.

Conclusion

This is often left in the air; the story stops when the girl is rescued or sent to bed in punishment. When the girl goes to hospital as a result of burns, she usually recovers; sometimes is met at home with chocolates and sweets; she is rescued from fire by fairies (a single instance); she leaves bad parents and goes to live in the hills (one instance); she promises never to light a fire again and is forgiven.

Picture 6. A man and a woman facing each other and gesturing, and a child sitting on a chair a little way off, watching them (designed to probe the child's reaction to a clash between the parents)

The couple are always seen as parents of the child sitting on the chair. They are nearly always described as 'arguing', or 'quarrelling'. The subject of 'argument' is often the child and his naughtiness, or his future ('what is to be done' about him or her). The family are often said to be 'poor' or 'very poor', and the parents are said to be discussing whether to send the child away to a Home, or to a rich relative where he would be properly fed. Other subjects of 'argument' or discussion between parents are as follows: father going away to a job far away to which mother objects; mother is pleading with father to get a job; parents are blaming one another because of the accident to the child; parents are too old for work and the child decided to earn money for them; the man is a drunkard and does not remember having lost his money; mother wants the father to buy clothes for the child and father says he has no money; *another* child is missing; they discuss the child's illness; a foundling child is found in the woods and adopted; father reproves mother and child for untidiness; there is no money on Christmas Eve; parents discuss what school the child is to go to; they talk about buying a new house; about Christmas presents for the child; about buying things for the house, such as washing machine; discuss having another child; discussing what to call the child; determining to leave the child in the woods because of naughtiness; discussing how to get more chairs because the child occupies their only one; father out of a job, no money for food or clothes; wife is asking for money; arguing as to whether the child is to be a nurse or a teacher; arguing as to whether the child can go on a school journey; playing a game of 'I spy'; arguing about husband's overspending; about who will have the child when they get divorced on account of the child's naughtiness; discussing if the child can go out alone; whether the spots on his face mean that he is ill; discussing a move to another house; father wants to join the Army; arguing who would use the only chair; wife asking for money to buy new clothes or a new hat.

Child

(*a*) '*Naughty*'. Shouts at father; puts knees on chair and spoils it; throws stones at pet dog; knocks over a vase; buys a lot of toys when sent out to buy food; will not take his medicine; runs away because he does not like the parents; smashes his own windows; makes holes in chairs by climbing over them; is not nice to visiting friends, throws a girl's doll into the fire; truants from school; spends dinner money on sweets; would not sit down properly but always kneels on chairs; breaks several chairs; plays about on furniture; tears up chair covers; throws his shoes away; upsets a pot of porridge on himself; breaks glass in a green-house; does badly in the 'exams'; breaks a lot of cups; steals mother's necklace to get the father into trouble because he broke her doll; jumps on the bed and makes a mess of the sheets; usurps father's chair.

(*b*) *Good*. Helps father to find a job; finds a treasure and makes the family rich; finds the missing object about which the parents are arguing; helps to find the missing baby sister; is unwell; is a bit frightened; is wondering whether the parents are talking about him; is hungry; gives his pocket money to buy food or help parents; forgives the parents who leave him and gives them money when grown-up and rich; gives mother the food he had saved when pinching from the larder; runs away because he is hurt by parental rejection; does not like the parents arguing; helps them when grown-up despite rejection in childhood; puts out an extra effort at school; makes chairs for all the family; wants mother to have new clothes or hat; promises to be good and keeps his promise; tells parents not to quarrel.

Parents

(*a*) *Love*. Parents forgive the child for naughtiness; persuade the child to take medicine and make him better; take the child to hospital; are worried about the child being ill; call the doctor to see the child; change their mind about sending the child away on account of naughtiness when they find him crying; ask the child what presents he wants; mother begs the father to buy the child a train-set; mother stops father from striking the child, explains the breakage was an accident; father agrees to keep the child on mother's request; father refuses to spank the child as mother wants him to; discuss child's birthday present.

(*b*) *Reject*. The child is put to bed as punishment; parents do not like the child, but bring him back home when he runs away; smack the child; let friends take her away and keep her; lock the child out of the house; leave the child in the woods; send the child away to a Home for three years; mother is telling the father how naughty the child had been; wants him punished.

(*c*) *Disagree*. Quarrelling about the child whom one of the parents had punished unfairly; mother stops father from smacking the child; parents are arguing whether the child is to be a nurse or a teacher. Father is usually represented as the harsher of the two parents, but occasionally it is the mother who insists that the father spank the child and the father sometimes refuses; it is usually the father who thinks the child must go away because of bad behaviour and the mother pleads, and sometimes prevails on him, to keep the child at home. Parents argue as to whether the child should be allowed to have the only chair; they argue also about who should take the child on holiday. In one case the mother wants 'to get rid' of the child, but the father wants to keep him; sometimes this situation is reversed; selfish mother wants clothes for herself and kindly father in league with child arranges for a school journey instead; mother slaps father for slapping boy, they argue how the child should be punished.

Conclusion

Child runs away but is found; parents repent and live happily reunited with child; child goes away from rejecting parents and likes it (rejects parents); the family's financial difficulties are solved sometimes with the child's help (work or money); 'next morning all is forgiven' or 'forgotten'; 'naughty' boy reforms after a police reprimand; father wins £500,000 on the pools, father finds a good job; boy goes to a good school; parents give in to the boy's habit of kneeling on chairs; child is rewarded for doing well at school; child reforms after three years in a Home; father takes mother and child out on mother's insistence; child falls ill with polio and is cured; a rich aunt supplies food to the poor family and the child is not sent away; child reconciles parents; lost money is recovered and the family get a country home; lost purse is found; the parents forgive the child's naughtiness; child returns home from a rich aunt with clothes and money; child finds a cache of gold; a benefactress (rich lady) buys a house for the family and employs mother; poor parents buy a train-set for sick child who recovers.

When the child is 'naughty' the story sometimes ends simply by the parents sending the child to bed.

COMMENT

Only very few of the 200 children tested showed originality of invention, or what could be described as a gift for composition; nevertheless 96 sets of 7 stories each, i.e. 48 % of the whole, could be classified as 'imaginative and lively'; 68 sets (34 %) as 'realistic, simple and factual in detail'; 13 sets ($6\frac{1}{2}$ %) as tentative (hedged in with 'could be', 'might', 'I suppose' 'it looks like'); and 56 sets (28 %) as 'revealing preoccupation with a specific theme or subject'. Of this latter group 6 sets (3 %) were concerned with clothes of the characters, their names or ages; 10 sets (5 %) with accidents, illness, injections and hospitals; another 10 sets (5 %) with crime; 6 sets (3 %) with adventure or feats of courage and resourcefulness, and 24 sets (12 %) with fairy-tale themes and magic.

Some of these categories overlap, as many stories dealing with crime, adventure, accidents, or having fairy-tale subjects, were also classified as 'imaginative, lively', and some of the 'tentative' or revealing preoccupation with specific themes, such as names, age or clothes, came under the heading of 'realistic, simple, factual in detail'.

About 8 % of the total number of stories gave indication of emotional disturbance and unsatisfactory inter-personal relationships because of the prevalence of sado-masochistic themes in them. Such stories occurred among the 'imaginative, lively', the 'fairy-tale themes', the 'accidents' and the 'crime' categories. These presumably disturbed children may have been 'slipped in' among the 'normal' because their disturbance had remained undetected and they passed as 'normal' in school.

CONCLUSIONS

The analysis of the responses of 200 normal children between the ages of 8 and 12 did not bring out any new patterns as compared with those that had been established in the earlier investigation. These patterns are predominantly positive in character and based on the reality of the child's everyday life. By looking at the data of Table 1, an investigator of Family Attitudes could feel fairly confident of what type of response he could expect from a normal child; and accordingly to detect deviations from the pattern which would indicate the desirability of a further investigation of the child's emotional condition.

Deviations from the 'normal' pattern cannot of course be expected in every story, and in some may be only slight, but the general trend in the stories of the disturbed could be expected to show a harsher picture of reality, a greater wealth of sadomasochistic fantasies and of 'unhappy' endings than shown by the 'normal'. On the other hand, the stories which are unrealistic because they avoid any mention of conflict and negative feelings could be regarded as deviating from the 'normal' on that account.

(Two examples of sets of stories by severely disturbed children are given in the Appendix, and two of stories by normal children, matched as closely as possible with the disturbed in age and level of intelligence.)

APPENDIX

Examples of stories by disturbed children

Girl, aged 9 years 10 months
Intelligence Quotient, 124
Siblings. 1 brother, aged 12
Diagnosis. Anxiety state with underlying depression

This is a set of stories in which hostility, guilt, depression and anxiety are expressed with an unusual consistency without a single relieving trait. In this girl's case, the mother suffered from severe depression during the child's babyhood; she was in charge of a neighbour from the age of 3 years while mother went to work. The father was an unstable person subject to irritability and bad temper when he 'let fly' at the girl. She was jealous of her elder brother and deeply resented her mother's going out to work throughout her childhood.

Picture 0

There was a woman. She wanted a baby. She got one and so she bought a nice cot for her.

(Q) Baby is sometimes naughty, sometimes bad.

(Q) She hits mother and all sorts.

(Q) Mother goes and tells husband. (Q) He tells baby off.

(Q) It's a girl.

(Q) She likes the baby a bit but not as much as she could do. (Q) If she asked husband to hit the baby she can't love it properly.

Picture 1

There was this Mother. She had a husband and a big son and a little child—they had quite a nice home. The mother was quite fond of the boy and of the boy baby too. The big brother grew older and got married. He asked his mother for money and she gave him some money and he was very foolish because he spent it on a wedding ring for a horrid wife. (Q) She was horrid because she didn't really love him. She only made out she did. He found out she wasn't nice and he divorced her and went away to far off and she (the mother) did not see him any more. Not all stories end happily. I would not leave my mother like that.

(Q) The little boy is a good boy.

(Q) Mother and Father like the little one.

(Q) The parents liked the big one till he went away and not now.

(Q) The big one doesn't like the little one—because he is always running after the big one. He caused an accident. He tripped the big one and he nearly fell in the fire.

(Q) Oh yes, the little boy likes the big boy.

Picture 2A

This mother and father had a baby. The girl was very jealous and when mother played with the baby the girl pushed Mother and Father away, and as she grew older she found she was very stupid but couldn't stop because she'd been doing it since a baby. She's tried to kill the baby. Mother got a cat for baby—the girl strangled it and said it was the baby. She did the same to the dog. Mother found out later when the baby talked. The mother put her (the girl) in an orphanage and she died and she felt she'd died of sin because she did all that to the baby. Normally, before the baby, she was good—I would not call her good now.

Picture 3

Ooh—There's a little boy. He can't walk. He's a cripple. Mother and Father didn't like him. He sat down for 4 years. He tried to walk. He couldn't. The ambulance came to pick him up. But the hospital couldn't help. He couldn't walk. Went back home and sat in the chair. In bed he was still in the sitting position with his legs up and he's all cold.

(Q) The Mother and Father are talking behind his back about the horrid things they'd do to him. They'd put him in the fire and leave him in it. He'd be helpless and he'd burn.

(Q) He's thinking if only I could walk. (Q) I'd go and tell Mother and Father—they'd be unhappy about him walking.

(Q) So send him to a very horrid Home and likely he'd run away, so if not right he'd go into hospital.

Picture 4

Ooh—Once a poor woman had a baby and she couldn't give it much clothes. A bully of a Father. Mother loved baby. Father said if she wouldn't let him punch baby he'd kill her. So she let him. Baby had red marks on its face. She went to hospital. Mother was hurt too. Father had murdered a bank manager, found guilty and was hung. Mother was happy but unhappy too. If only Father had not been so horrible—but no good thinking—the past is the past.

Picture 5

This Father is very like the other father. He had this child. She grew to love insects and go out in the morning and look at them and say how unhappy if they were cut down. Father got furious—he hated nature. He got her a pet hamster. Father let it go. It had no food. In the end it died. Girl said 'Why did you do it?'. 'I like to see animals suffer.' She cried. Mother got her another in the end. No good—he let that out too. In the end she let last animal suffer. Girl went to see pets' home. Was allowed to bring back an animal each day and take it back next morning. She loved puppies. Father said she could keep it but Father killed it. Father was put in prison for almost murdering puppy.

Picture 6

Oh—Mother and Father had a boy, didn't really want him. They didn't know what to do. Thought they would send him to a Home. He made a lot of friends there. When it was time to go to his own home he was unhappy. Mother and Father forced him to go. He cried—In the end he went back to the Home but all his friends were gone. In the end he did go back home. In school he found his friends. His Mother and Father moved to hurt his feelings. Again he found his friends. His Father and Mother knew nothing. Mother and Father secretly dressed up to conceal. When they found him with his friends they moved and moved. In the end they put the boy in an orphanage and stayed at home. They died of hunger. At the funeral the boy was glad they had died because they were wicked parents.

Boy, aged 8 years 5 months
Intelligence Quotient, 95
Siblings. One brother aged 5 years 2 months
Diagnosis. Anxiety state with obsessional, compulsive symptoms

There is plenty of fantasy in this set of stories. The one about picture 0, as well as that about picture 2, were most likely inspired by the hostility towards the younger brother, as this boy was said to be very jealous of him. In the story about picture 2 the situation represented in the picture is ignored and an incident of kidnapping introduced instead. Jealousy and aggression towards the baby, however, is ascribed to the child in picture 3. In the story about picture 4 the suggestion of hostility is evaded and a story of a family party produced instead.

Picture 0 (no hesitation in starting and practically no delay)

Once upon a time a woman lived in a cottage and she had a baby and was very poor. And the father was rich, but he wouldn't let mother have any money. And one day he made up his mind to go away because the baby was always crying. And when he went away he made up his mind to find a job, and he was looking around to find a job. And he came to a barber's shop and this barber was very angry because he didn't like his job. And he said: 'I'll take it over for you.' And the years went by and the mother and the baby.... He grew up to ten, and she made up her mind...he had a lot of hair, and she went to take the boy to the barber's, and it happened that she went to the barber's shop that the boy's father was working at, and the father did not notice the boy because he's only seen him when he was a baby. And after the barber cut his hair, it was done so nice that she made up her mind to go there every time. And one day the barber noticed her face and asked: 'What's your name?' She said: 'Mrs Brown.' And he asked: 'Where do you live?' She said: 'I live in a cottage which is near a lot of other cottages.' And she said: 'And I live at the second house along the cottages.' And the barber says: 'I am your wife, Mr Brown.' Mrs Brown said: 'Why did you have to go away.' And she (he?) said: 'Well, now your baby has grown up I shall leave this barber's shop and return to the cottage with you.' And then they lived happily ever after.

(Q) ('I suppose the father didn't like the baby then?'). 'No, he didn't.'

(Q) ('But the mother liked him?') 'She had to....'

(Q) ('She did like him?') 'She had to like it, she couldn't leave it; it would die....'

(Q) ('Was he a nice baby?') 'He was nice.'

Picture 1

Once upon a time there lived mother, father and a little boy. One day the boy got a smack because he's been told to deliver a message to his auntie and the boy didn't do it. And the mother didn't find out...the mother found out and told the father. That's how the boy got the smack. And two weeks later the father said to the little boy: 'It is your birthday today and you are six.' And he said: 'Your uncle's coming to tea and he might bring you a present. So be a good boy.' And the boy said: 'All right, Daddy.' The afternoon came and the little boy was very happy, and at four o'clock a knock came at the door. And the father said to Johnny: 'Open the door, Johnny', and he says: 'All right, Daddy.' As the boy was running to open the door, he said: 'That might be Uncle.' He opened the door and there was uncle! His uncle said: 'Happy birthday, Johnny! You're six years old.' And he says: 'Come on, Uncle, come on in!' And uncle came in, and they had all the tea ready and mother and father said: 'Hullo, Uncle! We've got time.' And Johnny was never naughty any more, and he was very happy as well.

(Q) ('What about the present from his uncle?') 'He had that as well. He got him a garage with three lovely super-cars, and he played with them and lived happy ever after. (Q) ('Did Johnny like his uncle?'). 'Yes, he liked him.' (Q) ('What about the uncle?') 'Yes, and he liked them.'

Picture 2

Once upon a time there was a woman called Mrs Jones, and she decided to have a baby, and she asked her husband would he agree. He said: 'Yes, all right.' Then: 'But remember we have to make up a name for him.' When they got the baby, Father came home from work and said: 'Have you got the baby?' 'Yes, yes' (excitedly), 'come and see him. He's in the bedroom.' The father rushed into the bedroom quickly and on the bed laid—nothing. And the father ran out again and said: 'The baby's gone, come, 'phone the police!' And she said: 'All right then' (rather casually). She dashed to the phone, she got the police and the police came round and said: 'Have you looked?' 'Well', she said, 'we went into the bedroom, we expected to see the baby there, but it was gone.' They said: 'Search the room and the bedroom.' So two men ran into the bedroom and two men looked all over the room and couldn't find anything. So the officer said: 'Go outside, he can't be far.' So two men went out and looked round as he might have run out of the door. Suddenly they saw a man run down the road and they saw a baby;

and one of the men said: 'Well, there's a man running down the road with a baby. What does he look like?' Father said: 'Curly hair and...and he had a pair of socks and no shoes.' And the policeman said: 'That's him.' So the policeman dashed out again and he saw no sight. And just round the corner he saw a person running round a corner. So he said to the other copper: 'Come on, Cliff', and then they both jumped into a police car, dashed round the corner and there they saw a man. They dashed right up to the top of the road and they both jumped out and hid. Soon the man came closer and closer and just as he passed, they ran out and they jumped on him. And one took the baby and dashed down the road in the car and gave the baby back. And the officer said: 'You should be well rewarded.' As this was going on, the other policeman held the robber and walked him to the police station, and he was put in prison for...(a long pause. Ther.: 'For how long?') Eight years. And so mother and father thanked both men and the officer and the three men walked out and the mother and father lived happily ever after.

(Q) 'You haven't said anything about the bigger child—the boy in the picture.' (After a slight pause) They got another baby in six years' time. (Q) 'And what is the elder boy thinking of the baby?' 'He thinks he's being brave and good.' (Q) 'Brave? What has he got to be brave about? Kidnapping?' 'No, they keep them safely. They both are (safe).' (Q) Which do the parents like best, the baby or the bigger boy? 'Both. They like them both best; both the same.' (Q) What is the elder boy like? Is he good? 'Well, he's naughty sometimes, but when he's good, he's very good.'

Picture 3

Once upon a time there was a boy and he had his parents. His Dad was called Jim and his mother was called Mary. And one day when they were going out they decided to have a baby, and he said 'All right then' (Who said this? The boy?). 'No, the father.' And they went to the hospital and they were told to come next week, on Wednesday, at 10.30. And they got the baby and the mother and the baby were all well. When they got home, the father said: 'It's very nice, the baby.' And one afternoon when television was on, he (the boy) saw some films, and afterwards he says: 'I like that film', and then he started to be naughty. And then he walked out and he went out to the baby and he got his bottle of milk and threw it on the floor, and then he hit the baby. And then he started to do other things naughty, and then his father came up to him and smacked him hard. Then he put him in his room, locked the door, and he said: 'You must stay there until it's your bedtime.' And then when it was his bedtime he let him out and sent him to bed, and in the morning, he said: 'I'm sorry, Dad for being a naughty boy yesterday and all last week', and his father said: 'All right then, I forgive you.' And then Dad and he went down the park and his father bought him an ice-cream, they had a game of cricket and football, and then he rolled in the grass and had lots of fun with his father. That night they came home and they said to each other funny things. When they got home they all had tea and watched television and that night he wasn't naughty. And all next week he wasn't. And then he went to bed and the next morning he was very happy and he stopped behaving naughtily for a long time. And then they lived happily ever after that and Richard grew up to be a big boy.

Picture 4

(Laughs. ½ min.) Once upon a time there was a happy family; there was the mother, Joan, and the father was called Jim and they had a baby called John. And the baby was only one year old. They had a house of their own, and the next day was Tuesday, and it was the mother's birthday, and she was thirty-one. And in the evening the neighbours had lots of presents to give her and father had a new dress to give her. They invited some of the neighbours in for tea and they seen television, they seen the Adventures of Tub Boat Annie, and it was very funny. And after that they seen other things and had their tea at the same time. And at nine o'clock the neighbours went, but two stayed until ten. When ten o'clock came they went, and they opened the presents, looked at them and packed them away, she said: 'It's time for you to go to bed, John', and she said 'Good night', and she kissed him. And then Dad came in and the mother said: 'That's enough now, Jim', and he says 'All right then, good night, John', and John went to sleep and it was quarter to 11, and then eleven o'clock striked and Jim and Joan went to bed, and from then on they lived happy ever after. (Q. concerning the raised fist—why? etc.) He was pretending the baby's done something; no, the baby wasn't frightened because he heard him say funny things.

Picture 5

Is that in the park? Once upon a time there was a family and there was five in the family; there was two girls, the mother and father and one boy. And one day on a Tuesday the father Joe said: 'It's my birthday next week. Have you collected any money to buy anything for my birthday? I have got a little surprise to give mum.' And the next week came, and in the afternoon one of the little girls went into the bedroom, and on the dressing table she found a watch, and she took it to her mother and she said: 'Look what I've found on the dressing table.' She says: 'Oh, don't touch that, that's Daddy's present.' And she says: 'Put it back', and as she was putting it back she forgot what her mother said, and she was so slow that she put it on the floor, and she was playing with it, and she heard that it wasn't going. And she quite forgot all about what her Mother said; and the carpet was next to the fire, and the boy came in, playing cowboys, and he charged round the room with a bit of rope, and he was twirling it about when all of a sudden it hit the watch on the carpet, and it flew up in the air, and it went right into the fire, and it started to burn. And just then the father came. And she says :'Your watch went into the fire!' And he says: 'Who did that?' and he says: 'Come here!' shaking his fist. And she says: 'I didn't do it. It was John; he was playing cowboys and he had a big rope and he knocked it into the fire.' He says: 'Where's John then?' And just then John came out of the bathroom. And he says: 'You naughty boy!' And he hit him hard. The mother knew what was going on and she nipped out and bought another watch, and she says : 'Here!' And he says: 'That's very kind of you!' And then day after day they forgot about what happened and they were happy.

Picture 6

(2 min.) Once upon a time there was a family; there was three in the family. There was mother and father, Jim and Margaret, and there was a little boy called George. One day Margaret was complaining about George wanting everything. And she said: 'I can't do anything about it' (the father said this), 'except that it is near his birthday. I don't know why he wants some toys when it's near his birthday, or is it sweets he wants?' 'He wants even Easter eggs, sweets and Easter eggs. He'll just have to wait until Easter comes' (said his mother). 'He can't have Easter eggs, just can't have Easter eggs.' And he cried and cried sitting on the chair. The days passed and they forgot about it. One Saturday morning he heard his father and mother talking about his birthday. Then he says: 'I want some toys.' The father says: 'Not again! Well, it's your birthday next week.' And then he still cried until Friday came....No, when Friday came his birthday was here, and Jim (Dad) said: 'All right, what do you want for your birthday?' And he says: 'I want a farm set.' And Jim says: 'All right.' And he brought him back a farm set, and there was four horses to it, there was three cows and there was another three cows. There was ten sheep and two ponies, and there was fields for the animals and there was a house for the farmer and his family. And he had a great game with them, and he never forgot the day his Dad got him the farm. And he had great fun with it and his father and mother were happy, and so was the boy.

Examples of stories by normal children

Girl, aged 10 years 1 month
Intelligence Quotient, assumed average
Siblings. Two sisters, aged 20 and 16 years

Picture 0

Once a baby was born. It cried a lot for a few days. The mother was very busy and had to keep stopping to watch it. The mother had to stay by it till it went to sleep...till it cried itself to sleep. She went off then and done some more of her work, but often the baby would wake up and start crying again. The mother thought that this would not happen; she thought that she would have a baby that would not cry. A week went by and baby did not cry so much. It would go to sleep and not wake up till the morning. The mother was pleased when the baby went to sleep quick; then every day her work would be finished. (Q iii) A little girl. (Q iv) It was naughty at first and then it got better. It would just cry and did not go to sleep very well. (Q vii) Just rocked the cot. (Q viii) Yes, I think very much, but I told you she did not like it very much when it kept crying and she could not get on with her work. (Q ix) Oh, yes, he liked the baby very much as well, but not when it cried sometime all night and he had to get up in the night and tell his wife.

Picture 1

A very naughty boy would not do as his mother told him. Mother and father had a talk about it. The father said that whenever he does wrong he must be smacked for doing it. Again he done a naughty thing by not going to school. His mother smacked him as she had been told to; then he would sulk. His mother and father watched the tele and did not mind about him. Another punishment was not to watch tele and go to bed very very early. When his mother said 'Bed', he would sit on the floor and not answer her, for he did not like going to bed. His brother picked him up and carried him to his bed. He was very annoyed at this. He would not go to sleep till very very late. When his father and mother came to bed, he would pretend he was asleep. (Q v) No, because he is such a naughty boy. (Q vi) They only like him very much when he is good but that hardly ever happens. (Q vii) No, because they smack him (with a smile). (Q viii) He only liked them when they said nice things about him and didn't smack him. (Q ix) No, because he had to miss the television to carry his brother to his bed. (Q x) No, he doesn't like his brother at all because he doesn't want to go to bed when he's told and his brother takes him.

Picture 2A

A little girl had a brother; it was very young. Mother and father didn't care for the little girl very much, but the little baby they liked most. She had to take herself up to bed; before her brother was born, her mother would take her up to bed. She wished this boy had never been born. This little girl was a very good girl and helped her mother. Even if she did help her mother, her mother would not think anything of her. ('What about her father?') Her father is just the same as her mother: he thought more of the baby than what he did of the girl. ('What happened in the end?') The mother and father did not know that one night when they were out the child packed up her little case and left. The child was never seen again, only by other people. ('Were they sorry or not?') Oh, they were sorry for what they've done and they tried to get her back, but she gave them one more chance. She said: 'If you like the baby more than me, I will go away and never come back.' ('And how did it all end?') Well, they loved the baby and her the same. They managed to like her just as much as they did the baby.

Picture 3

A little boy had a stool. He would sit on it and think of all different things. He would close the door of his bedroom and not go downstairs until he was called. Mother and father were worried about this because they wondered what he thought about. All he thought about was his school. He liked school very much, and sometimes, instead of just sitting on the stool, he would take some paper out and write. The mother and father loved this little boy very much indeed, and he loved them. He did not see his mother much in the day for he was at school most of the time. He was very young, about five. He liked school because he liked learning things. ('What did he write?') He liked writing stories. ('What kind of stories?') Stories that he imagined, and liked to be and liked to find out things.

Picture 4

A man was teaching his son how to box. He was a boxer himself and a very good one; he wanted his son to become a man like him. As he (the son) grew older, he could fight better. At his school he had lessons at boxing and he was the best of his class. His teacher did not know what his father was. When he was very much older, he took up boxing and made a living that way. (Q i) When he was young his mother would hold him up so that he could box. (Q ii) No, she did not mind at all because she wanted him as well to become a boxer.

Picture 5

A child was walking along in a wood when she saw some beautiful butterflies. She wanted to keep them, but she knew it was cruel. Her father and mother had told her not to keep butterflies because they die quickly in a jar. Every day she went to the wood to see whether her butterflies were there but she was not successful every day. Sometimes she could walk through the wood with her mother and father. Every day after breakfast she would go and look at the place in the wood where her butterflies were. Once she went and her butterflies were not there and she never saw them again. She was very sorry because she liked these butterflies. ('What is she doing in the picture?') She is looking at them and pointing to them. ('Who is this other person?') Her father, telling her to come back. ('Where is her mother?') This is one of the days when her mother had to stay indoors.

Picture 6

A girl liked playing offices on a chair, but when it came to dinner time and her mother wanted the chair, she would not let her have it. Her mother had to pick her off, but as soon as she turned her back, she jumped on the chair again. This was not at all nice for a child. The mother told the father. He said there is nothing to do but buy a new chair. He said: 'We'll have to let her have that one up in her room.' The mother disagreed with the father. One said the girl could have it, and the other said not. In the end the father won, the child could have the chair. (Q ii) Yes, but she doesn't think it's necessary to argue about it.

Boy, aged 8 years 7 months
Intelligence Quotient, assumed average
Siblings. One brother aged 10 years and 1 sister aged 1 year 9 months

Picture 0

One day when I was rocking my baby she closed her eyes and went to sleep. The birds was singing and chirping. (ii) The mother. (iv) Good. (v) Not always. (vi) Gets in a bad temper. (vii) (a longish pause) Smacks her. (viii) She likes her. (ix) The baby's father is a nice father. He goes to work four days a week. (Q) Cutting the steel.

Picture 1

The mother and the father of the child are sitting together and the baby is behind. The mother and father are watching the television and the baby is playing. The baby hears another man coming through the door. The man is his brother. (i) A good boy. (ii) Always. (v) Very much. (vi) Yes, always. (vii) Not too much. (viii) Not always. When they put him to bed too early he doesn't like them. (ix) Yes. (x) Yes.

Picture 2

The father is glad that he's got a baby girl and so is the mother, but the little boy is not because he wanted a brother. (i) He's thinking that his baby sister is all right: he might as well have one. (ii) Good (after a pause). (iii) Not always. (iv) Well (a pause) like when his mother says 'Go to bed' he starts arguing. (v) Get annoyed with him and smack him. (vii) Both of them. (Q vii repeated) The same.

Picture 3

The boy is thinking what he can do today. (Q) He thinks to play football, but he has no ball, so he asks his mother if he can buy a ball. His mother says 'Yes', so he goes and buys a ball. When he's got the ball he goes to play football. Before that he asks all his friends to come out and play with him. (Q) He's a good player and he's a good goalee.

Picture 4

(A longish pause) The mother is carrying her baby. The father is putting his fists up at the baby. He is very angry with the baby. The baby starts to cry. (Q) The father is angry because the baby cries all day. (Why?) Because he don't like going to bed. (ii) No. (Q) No, the father is not going to do anything (to the baby).

Picture 5

The little girl is making a fire and she is throwing on some sticks. The gardener comes along and tells her not to make fires in his garden. He tells her to put the fire out and go away. So she gets some earth and spreads it over the fire. (Q) Yes, she puts it out. (i) The girl's mother does not know she is there because she is in her house. (ii) Puts her to bed straight away.

Picture 6

The mother and father are arguing, and the baby is sitting on the chair listening. He does not know what they are talking about. (iii) He's thinking he'll go to his bedroom and play with his toys instead of staying there. (Q) They're arguing who is going to drive the car today. (iv) In the end the mother drives the car. (Q) No, the little boy doesn't mind who drives the car. His mother took him down to the park in the car. He took his bike.

REFERENCE

JACKSON, L. (1950). Emotional attitudes towards the family of normal, neurotic and delinquent children. *Brit. J. Psychol.* **41**, 35–51; 173–85.

(*Manuscript received* 11 *June* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 355–364
Printed in Great Britain

AN EXPERIMENTAL STUDY OF IMAGERY AND ITS RELATION TO ABILITIES AND INTERESTS

BY K. RAY CHOWDHURY AND P. E. VERNON
University of London, Birkbeck College and Institute of Education

Student subjects were assessed for four main types of imagery employed in carrying out short mental tasks, on the basis of their introspections. Some relation was found between these categories and regularity-irregularity of respiration, as claimed by Golla and Short; but no consistent relation was found with ordinary questionnaire self-reports nor with some of the alleged tests of imagery.

There was slight confirmation only of the supposed relevance of visual imagery to spatial ability. However, a number of associations at moderate to high levels of significance were found with (self-rating) tests of interests in different fields, and to a lesser extent with information about these fields.

I. INTRODUCTION

One of the areas in which individual differences were first studied was that of mental imagery, as a result of Galton's (1883) famous 'inquiries' with his 'breakfast-table questionary'. His finding that many scholars and scientists reported complete absence of visual imagery launched a long train of investigations into the functions of imagery in thinking. Bartlett (1927), for example, in a symposium in this *Journal*, held that imagery tends to appear when blockages occur in trains of thought, but that visual reproductions of experience are generally more distorted and condensed than verbalized memories. Many other writers supposed that individuals could be classified into imagery types—visiles, audiles, motiles, etc., or more simply into visualizers and verbalizers. It was assumed that these differences played an important part in the perception and retention of different kinds of information: that visualizers, for example, should be taught differently from verbalizers, and would be adept at different skills. Since, however, there seemed to be no reliable methods of establishing a person's type, and probably mixed types were commoner than pure ones, interest in this topic waned, and there has been little investigation in recent years.

Among the more systematic studies of introspective data were those of Betts, Carey and Burt. Betts (1909) expanded Galton's questionary so as to give a fair chance for images of every modality to be reported, and asked his subjects to grade each image on a 7-point scale from 'perfectly clear' to 'no image'. His results indicated that those who ranked their imagery high or low in one sense tended to do about the same in other senses; also that there was little or no relation between the use of visual imagery in various intellectual tasks and the subjects' abilities at these tasks.

Carey (1915) applied several of the alleged objective tests of imagery, but found these of little value. However, she claimed that reliable introspective reports could be obtained with school children, and noted that there was a tendency for negative correlations between strength of visual imagery and tests of higher mental processes, or *g*, suggesting that imagery might even be detrimental to academic studies. Again, in 1947, Brower found no relation between the reported intensities of images of various modalities and the Otis intelligence test. That questionnaire techniques may be made to yield more meaningful results by applying such complex types of analysis as between-person correlations and factorization was shown by Burt in 1938. Twelve subjects were selected from an earlier study by Burt and Davies who had been found to possess exceptionally vivid images of one particular type, whether visual, auditory or motor—4 subjects for each type. These subjects were given a questionnaire which required them to grade the strength of their mental images of about a hundred different experiences. The 12 sets of gradings were intercorrelated and four factors were extracted. The first or general factor presumably indicates the tendency which all persons have to recollect certain experiences more vividly than others, regardless of their own particular imagery and of the imagery suggested by each experience. The three remaining factors evidently represented visual, auditory and motor types of persons.

A number of more objective tests have been proposed for the assessment of strength of imagery, though in no case is there any convincing evidence of validity. Useful summaries are available in Carey's (1915) article and Woodworth's (1938) text-book. These include the Letter Square, based on the ability to reproduce letters arranged in a square matrix in a different order from that in which they were learnt (cf. Myers, 1911). Similarly, Fernald (1912) supposed that visualizers would be better able to spell words backwards from their visual images than would verbalizers. In the testimony or Aussage test, visualization is supposed to be a help to fullness, though not to accuracy, of description. Kraepelin is credited with using tests of fluency of association with visual and auditory stimulus words. Bowers (1931) found a small but significant correlation between the ease of learning visual and kinaesthetic words and the clearness of imagery claimed for them, but no corresponding correlation for auditory words. Roe's (1951) well-known studies of scientists indicated that imagery is relevant to choice of vocation, the more theoretical being more inclined to be verbalizers than the more experimental workers. She claimed that verbalists gave a much larger total of responses to the Rorschach inkblot test than did visualists.

Still other writers have used tests of spatial ability or *k* factor, assuming that these depend on visual imagery—for example Clarke (1937) and Ormiston (1939). The introspections of El Koussy's (1935) subjects indeed suggested that subjects frequently resort to visualization, but hardly justify the inference that success at such tests can be used as a measure of strength of imagery. A more productive approach is that of Barratt (1953), whose subjects rated the strength and other attributes of their imagery when performing spatial and other types of tests. He was able to show that those who rated their visual imagery higher on the space tests did score more highly, whereas no such differences occurred on non-verbal reasoning tests.

Quite a different attack has been through the study of respiratory rhythms, which appear to vary with the type of mental imagery. It has long been observed that there are individual differences in regularity of respiration, and Angell & Thompson (1899) considered that this depends on the degree to which the subject is in a state of equilibrium, undisturbed by environmental changes. They found that the rhythm was generally regular also during voluntary attention, but that it became irregular when there was a tendency to verbalize, as when the subjects were asked to recall nonsense syllables.

Golla & Antonovitch (1929) examined the regularity of breathing of 67 normal subjects, claiming an association between visual imagery and regular breathing on the one hand, and auditory imagery (with a tendency to vocalize) and irregular breathing on the other. Their findings were confirmed by Wittkower (1934) and Paterson (1935), who also compared the breathing records of psychotic patients and normals. In 1943 Golla, Hutton & Walter investigated a further physiological concomitant of imagery, namely the electrical activity of the posterior areas of the brain, recorded by means of the electroencephalogram. They distinguished three main types of alpha records: (i) M (minus) type with alpha rhythms below 10μV., on which opening or closing the eyes or mental activity had no effect; (ii) R (responsive) type with alpha rhythms within the normal range of 10–50μV. when the eyes are closed; this was blocked or much attenuated with eyes open, and diminished if mental activity proceeded while the eyes are closed; (iii) P (persistent) type with alpha rhythms of normal amplitude, persisting relatively unchanged during opening of the eyes or mental activity. They claimed a high correlation between M type and visualizing, the P, and to a lesser extent, the R type being more inclined to auditory-kinaesthetic or verbal imagery.

Short (1953) adopted the above methods with certain modifications. Instead of recording the respiratory rhythms with a pneumograph and ink-tracing device, he placed a thermocouple over the nostrils to convert the breath temperature into electrical energy for recording with the EEG amplifier. Short examined epileptic and other patients at the Burden Neurological Institute, together with 75 normal subjects, mostly university students. On the basis of introspections while carrying out a series of mental tasks, he classified them into two main imagery types, visual and verbal-motor. The visualists breathed regularly and showed frequent blocking of alpha-rhythms, while the verbalists breathed irregularly and showed alpha-persistence. However when verbalists attempted a problem requiring predominantly visual images, then physiological changes characteristic of visualists were apt to occur.

The present investigation was planned, first, to repeat Short's work in part and to compare his method of classifying imagery with the traditional questionnaire and alleged objective tests, since we felt considerable doubts as to the possibilities of categorizing individuals into consistent types. Secondly, we wished to explore the associations, if any, between individual differences in imagery and abilities or interests.

II. THE PRESENT INVESTIGATION

Subjects. Graduate teachers under training for the Postgraduate Certificate Course in the Department of Education, Kings College, were chosen as a suitable group of subjects for several reasons. First, it seemed desirable to employ intelligent adults, since stress was to be laid on detailed introspective data. Secondly, it was desirable to have subjects possessing a wide range of fairly stable interests, and this was likely to be true of students in training as teachers of languages, history, science subjects, etc.

Since the collection of introspections and breathing records involved individual testing, it was not possible to use as large a number of subjects as one would wish. However 34 men and 36 women (aged 21–26) eventually completed all tests between

January and March 1955. Good rapport with the students was obtained by a preliminary meeting, at which the objects of the investigation were explained, also by securing the help of student representatives in planning individual sessions.

Methods: (*a*) *Introspections and breathing records.* Practical difficulties made it impossible to obtain EEG records, and the breathing records were obtained with quite simple devices: (i) a clockwork kymograph giving one revolution in 30–180 sec.; (ii) a rubber sack or pneumograph connected to: (iii) a Marey's tambour consisting of a shallow metal vessel enclosed by a thin rubber membrane, with a light lever resting on it; (iv) a stylus with an ink tracing capsule; (v) a Jacquet chronoscope.

After preliminary conversation and instructions, allowing about 4 min. for the subject's breathing to settle down, a 1-minute practice trial was carried out, consisting of the mental rehearsal of any story the subject could recollect and the calculation of a simple multiplication problem. The subject then relaxed with eyes closed for 5 min., and his normal respiratory curves were recorded for 2 min. Six tasks were then given in turn: a sum, recalling the story of Noah's ark, solving a 3-dimensional problem based on cubes, rehearsing the National Anthem, an argument and the Lord's Prayer. These lasted about 20 sec. each, 60 sec. for the cube problem. Each was followed by some 1½ min. of introspection on the mental processes used in carrying out the task, and a further 1 min. rest-pause before the next task. The total time per subject was about 30 min.

The remaining tests were given in two group sessions with the whole body of students, each session lasting about 2 hr. There was no time limit for the Questionnaire or for the Interests Survey.

(*b*) *Questionnaire.* This was based on a version devised by Prof. Burt. It asked for self-assessments of the vividness of visual, auditory or kinaesthetic imagery in 28 imaginary situations, and yielded three scores for the contrasted imagery types.

(*c*) *Fluency tests.* Subjects wrote down for 1½ min. each as many red objects, blue objects, harmonious or pleasant sounds, and harsh or unpleasant sounds as possible. These provided scores for visual and auditory 'fluency' and, when combined, a measure of verbal fluency. Since the mean correlation between two visual, or two auditory, lists was 0·599, and between a visual and an auditory list 0·502, it seems probable that scores depend much more on verbal ability than on the particular sense, or imagery, modality.

(*d*) *Rorschach test.* Following Roe (1951) and Golla *et al.* (1943), five plates from the Rorschach test (nos. 1, 3, 6, 8, 10) were shown on colour slides for 4 min. each, and subjects wrote any responses that occurred to them on duplicated outlines of the blots. These were not scored fully, but animal, non-animal and movement and total responses were counted; the proportion of non-animal to total responses appeared to be the most promising indicator of imagery type.

(*e*) *Ability tests.* The abilities thought to be most relevant were verbal and spatial, and the following tests were selected, most of which had been designed for use with high-grade personnel in the Royal Navy: (i) vocabulary test, SP 193–4 (10 min.); (ii) non-verbal *g* test, SP 192 (20 min.); (iii) copying (spatial) test, SP 97 (10 min.); (iv) paper formboard, SP 76 (10 min.); (v) English usage test (10 min.); (vi) examination marks in Principles of Education.

(*f*) *Interests survey and information tests.* An interests test was constructed along the lines of Strong's Vocational Interest Blank, with lists of occupations, school subjects, periodicals, leisure-time activities, etc., each item to be checked + +, +, 0, − or − − according to liking or disliking. The items were classified by content (not, as in the Strong Blank, by empirical group differences) to provide measures of the following interests: (i) artistic, subclassified into active and passive interests; (ii) business; (iii) mechanical-constructional (for men); (iv) domestic (for women); (v) gregarious-social; (vi) handicraft; (vii) literary, with active and passive subgroups; (viii) musical, active and passive; (ix) outdoor-athletic; (x) scientific; (xi) dramatic-theatrical; (xii) welfare-humanitarian.

In addition, a general knowledge test was constructed containing 12 corresponding sections (but not differentiating between active and passive artistic interests), each section with 10 to 16 multiple-choice items of information relevant to the interest. The total score on all sections of this test was taken as an additional measure of verbal ($g+v$) ability.

III. ANALYSIS OF RESULTS

Since some of the distributions of scores were skewed or irregular, they were all converted to a standard scale ranging from +5 to −5 for correlational purposes. In the case of interests and information tests, where significant sex differences were noted, the distributions were normalized separately, so that each sex obtained a mean of zero on each variable. No appreciable sex difference occurred in the imagery data.

Factorization of abilities. Table 1 shows the inter-correlations of eight ability tests. It will be seen that, in this highly selected student group, there is no trace of a general factor, and that the non-verbal test, designed as a measure of g, correlates strongly with the two spatial tests. A centroid factor analysis confirmed that the first three tests could be used as measures of spatial, and nos. 4, 7 and 8 as measures of verbal, abilities.

Table 1. *Inter-correlations of ability tests*

	1	2	3	4	5	6	7	8
1. Non-verbal g	—	0·532	0·669	−0·374	−0·014	−0·017	−0·335	−0·019
2. Paper formboard	—	—	0·655	−0·289	0·063	−0·151	−0·315	0·058
3. Copying (spatial)	—	—	—	−0·281	0·163	−0·046	−0·223	0·150
4. Vocabulary	—	—	—	—	0·130	0·144	0·504	0·567
5. English usage	—	—	—	—	—	−0·015	0·115	0·177
6. Verbal fluency	—	—	—	—	—	—	0·185	0·077
7. Examination marks	—	—	—	—	—	—	—	0·285
8. General knowledge	—	—	—	—	—	—	—	—

In an analysis of the reliability and consistency of the interests and information tests, the latter had been tried out on a sample of graduate teachers under training at Goldsmiths' College (14 men, 29 women). The mean reliability of the 12 subtests was 0·655. For the main group of subjects, the correlations between the two measures of each interest are shown in Table 2.

Considering that the tests were in an early stage of development, the amount of positive overlapping between these two approaches to interest measurement is fairly promising, though obviously rather irregular. We might, however, expect more consistent agreement between each student's pattern or profile of scores on the two

Table 2. *Correlations between interests survey and information subtests*

	Men	Women		Men	Women
Artistic	0·446	0·481	Literary	0·660	0·529
Business	0·285	0·123	Musical	0·649	0·197
Constructional	0·546	—	Outdoor	0·079	0·079
Domestic	—	0·083	Scientific	0·467	0·417
Gregarious	0·203	0·241	Theatrical	0·500	0·438
Handicraft	−0·099	0·262	Welfare	0·104	0·307

measures, and the following formula (based on a technique described by Allport & Vernon, 1931) was applied to the normalized scores:

$$R = \frac{\sum_{A}^{N} (X+Y)^2 - \frac{1}{n}[\Sigma(X_a+Y_a)^2 + \ldots + \Sigma(X_N+Y_N)^2]}{\sum_{A}^{N} X^2 + \sum_{A}^{N} Y^2 - \frac{1}{n}[\Sigma X_A^2 + \ldots + \Sigma X_N^2 + \Sigma Y_A^2 \ldots \Sigma Y_N^2]} - 1,$$

$A \ldots N$ = subjects; n = number of interest categories; X = score on each interest category; Y = score on each information subtest. The standard deviations for all scores are approximately 2 and the means zero.

The over-all correlations are still not very high, though better than the average coefficients for separate categories. It would appear that men show greater consistency between their interests and information than women do.

Table 3. *Correlations between interests and information tests*

	Mean of separate coefficients	Pattern coefficients
Men	0·349	0·480
Women	0·231	0·274

In scoring the breathing records, the distances between successive crests were measured in millimetres, both for all the rest periods combined and for all the task periods combined. From these measurements the length of each subject's median breath, together with his 90th and 10th percentiles were calculated. The two latter quantities provided a measure of the range or degree of irregularity of breathing span under rest, or under task, conditions. Since, however, a considerable correlation was observed between median and range (i.e. slower breathers tended to exhibit more variability), a correction was applied, based on the regression of range on median.

The correlation between the regularity scores at rest and during tasks was 0·240, which is significant only at the 5% level, suggesting that the characteristics of breathing vary considerably with the situation. However, the two measures were averaged in exploring the relations between the questionnaire and objective tests of imagery, shown in Table 4.

There is some confirmation for Short's claim that visualizers breathe more regularly ($r + 0{\cdot}378$), and possibly for a relation between Rorschach non-animal responses and kinaesthetic imagery. But the remaining coefficients negate any consistent relationship between different types of assessment. Thus the auditory questionnaire bears no relation to auditory fluency, nor the kinaesthetic questionnaire to breathing

Table 4. *Correlations between measures of imagery*

	1	2	3	4	5	6	7
1. Questionnaire: visual	—	0·368	−0·292	0·229	0·088	0·101	0·378
2. Questionnaire: auditory	—	—	−0·003	0·136	−0·024	0·031	0·167
3. Questionnaire: kinaesthetic	—	—	—	0·130	0·055	0·334	0·004
4. Visual fluency	—	—	—	—	0·624	0·410	0·106
5. Auditory fluency	—	—	—	—	—	0·394	0·048
6. Rorschach non-animal	—	—	—	—	—	—	0·218
7. Breathing regularity	—	—	—	—	—	—	—

irregularity; and the Rorschach score overlaps more closely with both measures of fluency.

Short has commented on the extreme difficulties of classifying introspections regarding imagery. We did not attempt to impose any *a priori* categories on our records, but first classified under some eight headings which seemed to emerge from the material itself. The numbers of remarks by each subject falling under these headings were totalled and tetrachoric intercorrelations were calculated between the headings to see how far they overlapped. In this way, four reasonably distinct categories were reached, defined as follows:

I. *Visual-passive*, including reports such as 'pictured to myself', 'imagining myself seeing', and images of scenes, objects, people, etc., *not* in motion. In view of the positive correlation of this category with regular breathing, a small weighting was also given to the latter variable in totalling the scores.

II. *Auditory-passive*, including reports such as 'conscious of, or hearing one's self-produced speech', and audition of sounds, music or speech from other persons or things. A small weighting was also given to auditory scores in the questionnaire.

III. *Visual-active against Verbal*. These two subcategories were so highly correlated inversely, that it seemed preferable to accept them as a bipolar factor. Visual-active includes reports on self-activity, e.g. 'I was looking at the screen in a theatre', 'I looked at the cube and divided it in pieces', also moving scenes such as 'animals going in two by two'. Verbal was restricted to such statements as 'I said to myself', but excluded awareness of mouth or throat movements. The Rorschach score was also given a small, arbitrary weighting, since it showed some correlations with the introspective totals (cf. Chowdhury, 1956).

IV. *Kinaesthetic*, i.e. consciousness of movements or contractions of the tongue, eye muscles, hands, etc. Irregular breathing was also taken into account.

Finally the highest 15 students and the lowest 15 on each of these 'types' were picked out, and their scores on the various tests employed in the investigation were compared. Tables 5–8 list any group differences between highs and lows that reached various levels of significance, using the two-tailed *t*-test. *t*'s which are asterisked are to some extent inflated, since these variables were actually included in determining the imagery-type scores.

There is a definite relation between this type of visual imagery and regular breathing, though less close than that found by Short. However, the written questionnaire responses not only fail to correspond with the introspective records, but tend to contradict them. Though there is no consistent tendency for these visualizers to score above average on spatial tests, they are weak in vocabulary and slightly superior on a non-verbal 'g' test. Mechanical-constructional information and scientific interests show a more positive relation, and artistic, handicraft, theatrical interests and information slight negative relations.

Table 5. *Concomitants of high or low visual-passive imagery*

Highs superior on	t	P	Lows superior on	t	P
Regular breathing*	5·22	0·001	Vocabulary	2·62	0·05
Mechanical-constructional information	3·96	0·001	Theatrical information	2·16	0·05
Scientific interests	2·76	0·05	Visual imagery (questionnaire)	1·96	0·10
Business interests	1·95	0·10	Artistic information	1·82	0·10
Non-verbal 'g'	1·92	0·10	Domestic interests	1·74	0·10
Domestic information	1·88	0·10	Handicraft interests	1·72	0·10
			Artistic interests	1·54	0·20
			Gregarious interests	1·52	0·20
			Theatrical interests	1·49	0·20

Table 6. *Concomitants of high or low auditory-passive imagery*

Highs superior on	t	P	Lows superior on	t	P
Auditory imagery (questionnaire)*	4·05	0·01	Regular breathing	1·56	0·20
Literary information	2·25	0·05	Visual fluency	1·40	0·20
English usage	1·97	0·10	Auditory fluency	1·39	0·20
Literary-active interests	1·80	0·10	Welfare interests	1·32	0·20
Mechanical-constructional interests	1·65	0·20			
Principles of education	1·61	0·20			
Visual imagery (questionnaire)	1·57	0·20			
Theatrical information	1·53	0·20			
Musical information	1·36	0·20			
Rorschach animal responses	1·34	0·20			

Table 7. *Concomitants of high visual-active against verbal imagery*

High visual-active superior on	t	P	High verbal superior on	t	P
Rorschach non-animal responses*	3·57	0·001	Outdoor activities	1·57	0·20
Welfare interests	2·86	0·01	Business information	1·41	0·20
Kinaesthetic imagery (questionnaire)	2·21	0·05			
Rorschach animal responses	2·05	0·05			
Mechanical-constructional interests	1·86	0·10			
Literary-passive interests	1·35	0·20			

Table 8. *Concomitants of high or low kinaesthetic imagery*

Highs superior on	t	P	Lows superior on	t	P
Irregular breathing*	4·07	0·001	Scientific interests	1·26	0·03
Domestic interests	3·73	0·001	Business interests	1·33	0·02
Handicraft interests	3·69	0·001			
Visual imagery (questionnaire)	2·17	0·05			
Artistic-active interests	2·08	0·05			
Artistic interests	1·95	0·10			

* These t-values are inflated. See text.

The auditory group is rather consistently, though to a small extent, superior on several tests of 'v' factor, literary information and interest. There is also some relation to musical information and interest, though not at an acceptable level of significance.

The visual-active, as contrasted with the verbal, type shows few significant relationships, except possibly with certain interests and, unexpectedly, with kinaesthetic questionnaire responses. However, as posited, there is a closer relation with Rorschach non-animal than with animal responses.

The kinaesthetic type shows irregular breathing, but claims strong visual rather than kinaesthetic imagery in the questionnaire. These subjects appear to be definitely above average in handicraft and active artistic interests, and (among women) in domestic activities; but there is no indication of any relation to athletic inclinations.

In conclusion: a greater number of relationships, at moderate to good levels of significance, have been established between the introspective imagery categories and measures (self-ratings) of interests than with either information or ability tests, or other measures alleged to be diagnostic of imagery. However there is certainly room for further work in reaching a more adequate classification and assessment of introspections about imagery. We should suggest that more extensive records should be collected in a variety of situations in order to explore the consistency with which subjects employ different modalities (thus the discrepancy between our introspective and questionnaire data needs to be resolved). It should then be possible to select representatives of relatively extreme (though not, of course, of pure) imagery types and to study their patterns of interests with more reliable tests.

We are grateful for the assistance given by Prof. A. V. Judges and Mrs A. M. Blackwell, as well as for the co-operation of their students.

REFERENCES

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. Abnorm. Soc. Psychol.* **26**, 231–48.

ANGELL, J. R. & THOMPSON, H. B. (1899). The relations between certain organic processes and consciousness. *Psychol. Rev.* **6**, 32–69.

BARRATT, P. E. (1953). Imagery and thinking. *Austr. J. Psychol.* **5**, 154–64.

BARTLETT, F. C. (1927). The relevance of visual imagery to the process of thinking. *Brit. J. Psychol.* **18**, 23–9.

BETTS, G. H. (1909). The distributions and functions of mental imagery. *Teachers College Contrib. Educ.* no. 26.

BOWERS, H. (1931). Memory and mental imagery. *Brit. J. Psychol.* **21**, 271–82.

BROWER, D. (1947). The experimental study of imagery. I. The relation of imagery to intelligence. *J. Gen. Psychol.* **31**, 229–31.

BURT, C. (1938). Factor analysis by sub-matrices. *J. Psychol.* **6**, 339–75.

CAREY, N. (1915). Factors in the mental processes of school children. 1. Visual and auditory imagery. *Brit. J. Psychol.* **7**, 453–90.

CHOWDHURY, K. R. (1956). *An Experimental Study of Imagery and its Relations to Abilities and Interests.* Ph.D. thesis, University of London Library.

CLARKE, G. (1937). The range and nature of the factors in Perceptual Tests. Ph.D. thesis, University of London Library.

EL KOUSSY, A. A. H. (1935). An investigation into the factors in tests involving the visual perception of space. *Brit. J. Psychol.* (*Monogr. Suppl.*), no. 20.

FERNALD, M. R. (1912). The diagnosis of mental imagery. *Psychol. Monogr.* **14**, no. 58.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

GOLLA, F. L. & ANTONOVITCH, S. (1929). The relation of muscular tonus and the patellar reflex to mental work. *J. Ment. Sci.* **75**, 234–41.

GOLLA, F. L., HUTTON, E. L. & WALTER, W. G. (1943). The objective study of mental imagery. *J. Ment. Sci.* **89**, 216–22.

MYERS, C. S. (1911). *A Textbook of Experimental Psychology.* Cambridge University Press.

ORMISTON, M. (1939). The bearing of general and special abilities upon scholastic success at the beginning and end of a secondary school career. *Brit. J. Educ. Psychol.* **9**, 164–73, 213–23.

PATERSON, A. S. (1935). The respiratory rhythms in normal and psychotic subjects. *J. Neurol. Psychopath.* **16**, 36–53.

ROE, A. (1951). A study of imagery in research scientists. *J. Personality*, **19**, 459–70.

SHORT, P. L. (1953). The objective study of mental imagery. *Brit. J. Psychol.* **44**, 38–51.

WITTKOWER, E. (1934). Further studies in the respiration of psychotic patients. *J. Ment. Sci.* **80**, 692–704.

WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

(*Manuscript received* 13 *May* 1963)

Brit. J. Psychol. (1964), 55, 3, *pp.* 365–368
Printed in Great Britain

THE TITLES OF PAPERBACK BOOKS

By W. P. BROWN
Bedford College, London

Ratings on Semantic Differential scales were used to assess the 'connotative meaning' of a set of book titles. Each book selected had been reissued in a paperback edition bearing a title different from the original title. Ratings of both titles were collected and compared. In most cases, the paperback titles were rated lower than the original titles on Evaluation and/or higher on Potency. The change of title is seen as an attempt to increase the 'package appeal' of the book, perhaps by enhancing its capacity to hold the attention of the prospective purchaser.

I. Introduction

One of the many anxieties felt about the 'paperback revolution' has focused on the employment of what Todd (1961) calls 'packaging techniques' in the production of paperback books. She particularly cites 'the lurid jackets on which pneumatic females may be trying to sell hot stories about call-girls but may equally be "generating a sales impulse" for works by Zola, D. H. Lawrence, and Simone de Beauvoir', as a device which illustrates 'the contempt in which the purveyors of the mass media hold their customers' (p. 104). It should be noted that she is not here criticizing the content of many paperbacks, as others have done (Orwell, 1944; Hoggart, 1957), but is rather raising the issue of the *fittingness* of the pictorial cover to the book behind it. The question of fittingness is one which is, in principle, open to empirical investigation; one might induce people to read a book without seeing its cover, and subsequently to judge the appropriateness of the cover for the book. More practical suggestions arise when it is recalled that many paperbacks are reissues of books which have already sold in hard-cover format. If we may safely assume that the 'packaging' of the original edition was appropriate to the book itself, then a comparison of the external properties of the two editions might go some way toward answering the question about fittingness: to the extent that the paperback displays external characteristics different from those of the original, the former are likely to include features which are not appropriate to the book itself.

The external characteristics of a book, however, include more than the pictorial content of the cover stressed by Todd. The title of the book is another such characteristic, and may also be more or less appropriate to the book's contents. While the majority of books reissued as paperbacks bear the same title as the original edition, in a minority of cases a new title is substituted. Again, if we accept that the original title was an appropriate one for the book, then evidence of material differences between it and the paperback title will bear on the issue of the fittingness of the latter. A study involving only the titles of books has the practical advantage that they can be readily detached from the books to which they apply. But there is a further consideration: since changes of title are comparatively rare, when they do occur there are presumably strong reasons for the change. While no doubt these reasons vary from case to case, there may be an influential consideration common to them all; the publishers of paperbacks may have some positive conception of what is a suitable title for a paperback—a conception to which the original title (however appropriate to the book) may fail to measure up. (The choice of pictorial cover designs might be influenced in much the same way.) If this were so, then one would expect that all changes of title would be in some respect similar to one another. This possibility was investigated.

II. METHOD

The study reported here was carried out on the titles of eight books, all fiction, with paperback titles which were not the original titles (see Table 1). An attempt was made to discover whether the two titles for each book had characteristically different 'connotative meanings' from one another. Sixteen women undergraduates rated the sixteen titles on thirteen Semantic Differential scales, i.e. seven-point bipolar scales defined by a pair of opposed adjectives. The raters did not know that any relations held among the titles, and, while some said that certain of the titles were familiar, none claimed to have read any of the books named.

The scales used were drawn from the study by Osgood & Suci (1955), and included four to represent each of the three main factors reported by them, Evaluation (*pleasant–unpleasant, beautiful–ugly, kind–cruel, clean–dirty*), Potency (*strong–weak, large–small, hard–soft, heavy–light*), and Activity (*active–passive, fast–slow, sharp–dull, hot–cold*); the remaining scale was *rough–smooth.*

The expectation, largely intuitive, was that the paperback titles would be rated lower on Evaluation and higher on Potency and Activity (and consequently *rougher*) than the original titles.

III. RESULTS

Intercorrelations among the scales indicated that the four scales measuring each factor agreed with one another to an acceptable extent, with the exception of the *hot–cold* scale. This exception appeared to be due to denotative contamination (Osgood, 1962) arising from the presence of the words *hot, warm, flame* and *ginger* in certain titles.

With eight pairs of titles being rated on thirteen scales, there were 104 comparisons between paperback and original titles to be made. The sign test was used to determine whether there was significant inter-rater agreement about the direction of the difference between each title and its pair on each scale. In all, forty-eight of the 104 differences were significant at or beyond the 0·05 level by the sign test (most were much more significant than this). Of the forty-eight significant differences, forty-one were in the expected direction. The other seven significant differences were very restricted in distribution, referring either to the largely invalid *hot–cold* scale, or to comparisons relating to one particular book (*The Secret Ways/The Lost Frontier*) which showed five significant differences in the opposite direction to that expected.

In order to simplify the data for the purpose of making separate comparisons for each book, composite factor scores were computed by averaging the responses of each rater on the four Evaluation scales, the four Potency scales, and the three Activity scales (omitting *hot–cold*). Mean factor scores for each title were computed over all raters, and the differences between paired paperback and original titles in mean factor scores are shown in Table 1. Positive sign in the table indicates a difference in the expected direction, and the significance levels referred to derive from the application of the sign test to the factor scores. It can be seen that, for the first book, there are significant differences on Evaluation and Potency which go against expectation, and that for the second book there is no significant difference on any factor. The remaining six books all show at least partial confirmation of the expectations.

Table 1. *Mean differences between paperback and original titles on the composite factor scores*

Paperback title	Original title	Factors of connotative meaning		
		Evaluation	Potency	Activity
The Secret Ways	*The Lost Frontier*	−1·11*	−1·41**	−0·88
Scarecrow Lover	*Rapture in my Rags*	+0·16	+0·33	+0·27
Murder by Neglect	*Harriet*	+3·56**	+1·31	+0·48
Flame in the Streets	*Hot Summer Night*	+1·97**	+1·08	+1·77
The Intruder	*Line on Ginger*	+1·24**	+0·85*	−0·54
The Passionate Pilgrim	*Miss Nightingale's Girls*	+0·58	+1·81**	+0·79
All Hands on Deck	*Warm Bodies*	+0·95	+1·14**	+1·60
Something Wild	*Mary Ann*	+1·78**	+2·03**	+2·29**

* $P < 0{\cdot}05$, ** $P < 0{\cdot}01$, on a two-tailed sign test.

IV. Discussion

The outcome of the investigation, then, is that for seven of the eight books there are differences between the paperback and original titles to which the Semantic Differential scales are sensitive, and that in a majority of cases the paperback titles are given lower ratings on Evaluation and higher ratings on Potency than the original titles receive. That the publishers should go to the trouble of devising new titles in this way suggests that they may consider it appropriate for paperback books to carry an aura of low Evaluation and high Potency. Why they should consider this appropriate remains an open question. When Todd describes their attitude as one of contempt for their customers, she is attributing to the publishers the belief that only books with certain unsavoury characteristics could sell to a mass audience, implying that such a belief is false, and asserting that to hold such beliefs and attitudes is discreditable. The vigour of this indictment no doubt stems from the implications of a recourse to unpleasantness (low Evaluation) as a source of attractiveness, with the spectre of a cultural Gresham's law (unavailingly exorcised by Williams, 1958) hovering in the background.

More limited considerations follow from the recollection that the function of packaging is to attract and hold the attention of the prospective purchaser. Perhaps publishers assess the suitability of a title or pictorial cover for a paperback in terms of its effectiveness as an attention catcher. Certainly several of the factors known (Berlyne, 1960) to influence the direction of attention, such as colour, sensory intensity, and the representation of biologically significant stimuli (the 'pneumatic females'), are liberally employed in paperback cover designs. As for the titles, while one might expect that words with unpleasant connotations would tend to be avoided, there is evidence (on the latencies of very simple reactions to words: Lanier, 1941; Osgood, Suci & Tannenbaum, 1957, p. 158; Newbigging, 1961) which suggests that, once attended to, unpleasant words may be more 'arresting' than pleasant ones. It is possible, therefore, that the 'packaging' of paperback books, including changes in titles, may effectively contribute to achieving certain of the publishers' aims. Such effectiveness, however, even if convincingly demonstrated, would not rebut a charge of contempt for one's customers in the absence of agreement on the worth of the publishers' aims, the legitimacy (rather than the effectiveness) of their techniques, and the compatibility of techniques and aims. There remains ample room for controversy.

Finally, it may be noted that the connotative meanings assigned to the paperback 'package' are not far removed from those of the contents of such packages, as described by some influential critics of the genre. Thus Orwell (1944) formulated the general ethos of certain paperback novels as 'the cult of power...mixed up with a love of cruelty and wickedness for their own sakes' (p. 245), and Hoggart (1957) has recently provided a similar and somewhat fuller analysis in his discussion of 'sex-and-violence novels'. The dominant themes of their diagnosis—cruelty and wickedness, power and violence—correspond, at least roughly, to the sprigs of low Evaluation and high Potency which were grafted on to the books considered here, when their titles were changed for publication in paperback form.

REFERENCES

BERLYNE, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

HOGGART, R. (1957). *The Uses of Literacy*. London: Chatto and Windus.

LANIER, L. H. (1941). An experimental study of 'affective conflict'. *J. Psychol.* **11**, 199–217.

NEWBIGGING, P. L. (1961). The perceptual redintegration of words which differ in connotative meaning. *Canad. J. Psychol.* **15**, 133–42.

ORWELL, G. (1944). Raffles and Miss Blandish. Reprinted in *Collected Essays* (1961). London: Mercury Books.

OSGOOD, C. E. (1962). Studies on the generality of affective meaning systems. *Amer. Psychologist*, **17**, 10–28.

OSGOOD, C. E. & SUCI, G. J. (1955). Factor analysis of meaning. *J. Exp. Psychol.* **50**, 325–38.

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

TODD, J. (1961). *The Big Sell*. London: Lawrence and Wishart.

WILLIAMS, R. (1958). Culture is ordinary. In Mackenzie, N. (ed.), *Conviction*. London: MacGibbon and Kee.

(*Manuscript received* 12 *July* 1963)

Brit J. Psychol. (1964), 55, 3, *pp.* 369–370
Printed in Great Britain

NOTES ON COMMUNICATION IN HOSPITAL

BY G. W. H. LEYTHAM AND R. R. HETHERINGTON
University of Liverpool

Communication between staff and patients in hospital has been discussed in a Ministry of Health pamphlet (1963), and at special conferences in Oxford. Some initial inquiries into this topic have already been carried out by the Nuffield Research Unit at Liverpool (reports on various aspects of this work are available from the Department of Psychological Medicine, Liverpool University). One of the main tools in this kind of research is the interview, itself a problem in communication. These notes report an attempt to assess some of the possible sources of error, such as how the questions are phrased and by whom they are asked.

New admissions to two medical wards, one male and the other female, were interviewed by an academic psychologist at the beginning of each week, and subsequent admissions were seen by a clinical psychologist later in the week. Patients were interviewed individually and in privacy. A standard set of questions relating to their illness, worries, and communications with the staff were asked by each psychologist and verbatim responses were noted. Altogether, 111 patients (56 male; 55 female) were interviewed, ages ranged from 15 to 76 years, with a mean of 53·7.

Comparisons between the two interviewers' samples, containing sixty-six and forty-five patients respectively, showed that they did not differ significantly with respect to sex, age distribution, marital status, social class, religion, consultant in charge, diagnostic categories, length of illness, or period between admission and interview.

Responses were classified and analysed by two other members of the unit. Only two questions produced significant differences between the interviewers: 'What did you feel when you knew you would have to come into hospital?' ($\chi^2 = 7{\cdot}78, P < 0{\cdot}02$) and 'Have they told you how you are getting on?' ($\chi^2 = 8{\cdot}00, P < 0{\cdot}01$). A further break-down in terms of the sex of the patient showed that in both cases the discrepancy was attributable to the males interviewed by the clinical psychologist: nearly all of them had been favourably inclined towards coming into hospital, but hardly any of them thought they had been told how they were getting on.

A study of the verbatim responses made by male patients to the first question was helpful. A typical reply to the clinical psychologist was: 'Oh my word yes, it was a good thing.' To the academic psychologist it was: 'Oh I didn't mind, no.' Neither of these answers could have been to the set question ('What did you feel when you knew you would have to come into hospital?'), and on reflexion both interviewers recalled that if the patient hesitated they tended to add a rider: one adding something like 'Well, do you think it was a good thing?', and the other asking 'Did you mind coming in?' They had thus induced opposite sets, the former wondering whether the patients had been glad to come in, and the latter whether they had been sorry. A further difficulty with this question concerned the tendency of some patients to confuse their feelings about coming into hospital with their subsequent feelings once they were there. Again, attitudes towards leaving home, while they undoubtedly influence feelings about going into hospital, are not the same as attitudes towards the hospital as such: a mother is reluctant to leave her young children, but a lonely old woman is glad of the company on the hospital ward. Finally, some replies were difficult to classify: a response such as 'I didn't mind' could be 'favourable', 'neutral' or 'unfavourable' towards the hospital depending upon the way in which it was said.

It is apparent from the above results that the original question will have to be broken down into several simpler and more direct questions; tone of voice will have to be noted alongside verbatim responses; and an overall impression recorded at the time by ticking one of the coding categories.

In reply to the second question ('Have they told you how you are getting on?'), the men interviewed by the clinical psychologist said 'No, not really' or 'They're waiting for the tests to come through', while those seen by the academic psychologist made remarks such as 'The doctor came this morning and said "be patient"' and 'I was examined today and he said I was improving'. Following up these differences it was found that the latter group had been interviewed shortly after ward rounds, while the former group was seen in the morning, some time after the last medical examinations. Some of the answers suggested that merely being seen by the doctor during the first week in hospital gave patients the impression that they had been told how they were getting on, but that as time went by non-committal statements, such as 'You are doing fine', conveyed less and less information and changed in the patient's estimation from a medical opinion to a social pleasantry.

The importance of the time factor in hospital research was brought out still further when responses to another question, 'Have they told you what's wrong with you?', were assessed. All the patients seen were interviewed a week later by a psychiatrist, who also asked this question. At the first interview 41·4 % said they had been told, and a week later the figure was 58 %. This did not represent a simple addition, however, as over a quarter of those who originally said they had been told denied it at the second interview. Either they had reinterpreted the information they had received during the first week, or they had forgotten that they had been told.

To find out just how well patients do recall medical statements, a consultant was asked to write down exactly what he said to forty-seven out-patients. About an hour later they were asked if they could remember what the doctor had told them. In this short time the group as a whole had forgotten 37 % of the original statements. Certain things were remembered better than others, and various distortions occurred: all subsequent visits were recalled but nearly 60 % of the instructions telling the patient what to do were forgotten or wrongly recalled! The number of statements remembered correctly was found to be related to age (younger patients recalled fewer), anxiety level (those with an average level remembered more), and amount of information given by the doctor (the range was two to fifteen statements, and nobody receiving more than four statements remembered them all).

Good and effective communications in hospitals are obviously very important both for the efficiency of the staff and the well-being of the patients. Much further research is needed in this area, and psychologists—aware of the factors which influence interview responses and recall, and familiar with information theory—could make a valuable contribution to it.

(*Manuscript received* 3 *January* 1964)

PUBLICATIONS RECENTLY RECEIVED

Oscillation Experienced in the Perception of Figures. By KRISTIAN HOLT-HANSEN. Copenhagen: Munksgaard. 1962. (*Hist. Filos. Medd. Dan. Vid. Selsk.*, **39**, no. 7.) Pp. 48. Kr. 9.00.

Dr Holt-Hansen differentiates between three types of effect in which movement is experienced in the absence of any movement of a physical object: (i) the autokinetic phenomenon; (ii) where a figure is difficult to localize although as a physical object it is motionless; (iii) where parts of a figure appear to move in relation to other parts although the figure is stationary. This last phenomenon appears under conditions discovered by the author during his investigations.

In his first experiment subjects viewed an isosceles triangle containing one obtuse angle. The side opposite to the obtuse angle was then seen to curve outwards away from the angle and, with continued viewing, movement effects were seen, the line appearing to curve to different extents, or to appear straight. Movements of the line were cyclical in nature and the rate of oscillation was measured by the subjects tapping, or adjusting a tone so that it varied in pitch together with the changes seen in the figure.

Two main types of rhythm were found, the first, which was the more common, being a fluctuation with a periodicity around 0·7 cyc./sec., the second, a fluctuation with a periodicity from between 4 and 10 cyc./sec. The phenomenon was enhanced when vision was intermittent during stroboscopic illumination, the optimum effect arising with a stroboscopic frequency of about 10 cyc./sec. Under these conditions the periodicity around 0·7 cyc./sec. was not affected by the stroboscope frequency. In further experiments, when the base of the triangle was perceived via one eye, and the remaining parts of the figure via the other eye, movement was still seen. From this it was concluded that the mechanisms responsible for the phenomenon were central rather than peripheral.

Holt-Hansen's work is clearly not yet complete, but the experiments so far reported are a model of clarity and ingenuity, and considerably extend knowledge of the conditions under which these movement phenomena arise. They also provide techniques by means of which the answers may be provided to further questions such as:

(1) The oscillations have much in common with figure-ground and reversible perspective phenomena: are these latter phenomena affected in the same manner by stroboscopic illumination as oscillating autokinesis?

(2) Is the oscillation dependent on external stimulation or do central mechanisms operate independently of stimulation? For example, would tachistoscopic observations taken at random small time intervals reveal the 0·7 cyc./sec. oscillation when plotted against a linear time scale?

(3) It can be shown that reversible perspective and other phenomena are influenced by both the position of fixation of the eyes on the figure and, when the eyes remain fixated, by fluctuations of attention such that the subject attends to one part of the figure or the other. In his discussion Holt-Hansen considered eye-movements, but no acts of attention analogous to eye-movements. Could oscillations be affected by instructing the subject to attend to different parts of the figure?

C. E. M. HANSEL

Structural Linguistics and Human Communication. By BERTIL MALMBERG. Berlin: Springer. 1963. Pp. vii + 210. DM. 32.

To a psychologist, the title of this book holds out a promise which is not fulfilled; the subtitle is more communicative: 'An Introduction into the Mechanism of Language and the Methodology of Linguistics'. In fact, Malmberg, Professor of Phonetics in the University of Lund, concentrates mainly on structural methods in phonetics. In this field he is an authority, and reports interesting information collected from innumerable sources and relating to a score of languages. However, his aim is broader than this, his point of departure being 'the place of language in human relations' with linguistics to be understood as 'a science of communication'. With this in mind, Malmberg attempts to link linguistics with information theory, building, as it were, on Cherry, Miller and Wiener, with his own tools and materials. The result, alas, is not wholly successful. In the first place, the author seems to have been confused by the uses of information theory by electronic engineers and—among others—by psychologists. Thus he does not make an adequate distinction between first and subsequent orders of redundancy (p. 34) and the statistical concepts and general uses of the theory sit uneasily together. Then, from the possibility that the nervous system uses a binary code, he attempts to introduce the Jacobson–Halle theory of distinctive features as a possible perceptual code. Unfortunately, he continually uses as an illustration of a binary code, the Morse Code, which is of course at least a tertiary code (dot, dash and pause). This is by no means a trivial objection since, after an extensive treatment of the binary hypothesis, it emerges that the distinctive feature theory also requires triple choice-points, a point which is but briefly mentioned. Malmberg then mentions three authors with the comment that their results 'do not seem to confirm entirely the hypothesis under discussion' (p. 127) with no further details and no references! Thirdly, although the term 'redundant' is used frequently, and in novel situations, no mention is made of the relevant psychological literature, and the perceptual implications of the term are not fully realized (in both senses). For example, with two possible cues for a discrimination, the more powerful is termed distinctive and the other redundant (p. 93).

A final major criticism is that although Malberg rebukes his predecessors, and some of his contemporaries, for their lack of scientific method, his own quoted experiments seem somewhat inadequate. For example, in examining the acoustic correlates of 'accent' (i.e. 'stress' by some definitions: there is little agreement in the field as to the use or meaning of these terms and concepts) '...the original varying pitch (of recorded Swedish sentences) was replaced by a monotone...To my ear, any accent difference between the words under examination had disappeared. When I listened to the words in isolation I was no longer able to hear which accent they had. This result would seem to be sufficient proof that pitch variations are indispensible for the accent difference' (p. 105). Without seeming to pay slavish tribute to experimental design, large groups and impressive P-values, it seems to the reviewer that the experiment is totally inadequate, particularly since other authors (e.g. Fry) have found contrary results for English.

The fundamental problem in experimental phonetics and psycho-linguistics is to find correspondences or correlations between the different levels of description of an utterance; the linguistic, the articulatory, the acoustic and the perceptual. Although Malmberg sometimes oscillates uncomfortably between, say, perceptual and articulatory description and distinctions, he is fully aware of the difficulties involved in establishing relationships, however conditional, between the levels, and illustrates these difficulties cogently. What is apparent in the book is the struggle of phonetics to achieve scientific status; to find a metalanguage and a methodology, and Malmberg's treatment of the history of ideas is illuminating, aimed, it sometimes seems, directly at 'those linguists of traditional orientation who would like to make contact with the new trends in the study of language'. One occasionally feels the pull of history on the author's pen, but he manages to convey the fascination of historical changes in language and the influence of background on perception. If he does too often assume that linguistic realities are psychological realities, and use the terms 'identify', 'recognize' and 'perceive' in such a way that he sometimes has two or three ghosts in the machine, he is not alone in that. Flow diagrams of the process help to understand the uses of such words, but his is inadequate, not fully applied and seems to have feed-back loops in the wrong places.

The book is not easy to read, the index is inadequate, there is no author index (a vital necessity in a book of this kind) and the references are not easy to find. Nevertheless, it occupies a place in the literature which the reviewer has not seen filled previously. It should be read by workers in psycho-linguistics but critically and not as an introduction to the field. Miller, although at least five years out of date, might prove a good starting point, but there is room for several volumes in this area.

JOHN MORTON

Problemes de Psycho-Linguistique. Symposium de l'Association de psychologie scientifique de langue française. By J. de Ajuriaguerra, F. Bresson, P. Fraisse, B. Inhelder, P. Oléron and J. Piaget. Paris: Presses Universitaires de France. 1963. Pp. 219. F. 16+T.L.

This symposium consists, first of all, of a number of longer papers summarizing work done in a particular area, each paper being followed by a discussion, and secondly, of a number of shorter papers each reporting an individual study. It is agreeable to read, clear, perceptive and, with a few exceptions, sophisticated from the linguistic point of view—in fact, typically French, thus confirming national stereotypes! We should however take note of the fact that it is the French language rather than nationality that defines the participants. The biggest contingent, after Paris, comes from Geneva; indeed Geneva provides three to Paris's two of the five longer papers. The French papers follow the American model, in their layout, in their experiments, and in providing up-to-date summaries not only of experimental work in psycho-linguistics but also (Bresson's paper is excellent here) of the relevant information theory and structural linguistic theory approaches. The Genevese report only their own work, rich in classificatory detail but inelegant, and—a great deficiency—provide no bibliographies. Their theme, that the learning of logical operations is pre-linguistic though language may improve our mastery of them, is initiated by Piaget and supported by two interesting papers on the intellectual development of dysphasics, one by Ajuriaguerra and co-workers, the other by Inhelder and Siotis. On this central theme the symposiasts, contrasting themselves on occasion with the Russians, showed unanimity, strong support coming from Oléron, who contributed the fifth paper, excellent in itself but unfortunately saying nothing on his own work on deaf mutes.

The shorter papers, which take up less than a third of the book, are of a lower standard than the others. One defect revealed by the experimental studies is to draw conclusions about our use of language as a whole from experiments, investigating the relationship between variables such as frequency, familiarity, and perceptual threshold, which use very small samples of words chosen on the basis of a single feature, such as frequency. As far as method goes, psychology emphasizes experiment, linguistics description; psycho-linguistics should make them more nearly equal partners.

W. H. N. HOTOPF

Science, Perception and Reality. By Wilfrid Sellars. London: Routledge and Kegan Paul. 1963. Pp. viii+366. 50*s*.

Philosophy and Scientific Realism. By J. J. C. Smart. London: Routledge and Kegan Paul. 1963. Pp. viii+160. 25*s*.

Both these recent additions to the International Library of Philosophy and Scientific Method represent a revolt against the still widespread view of philosophy as concerned exclusively with clearing up logical muddles arising from the misuse of language. In its place they suggest that it is the task of the philosopher to help us to reach an intelligible and comprehensive understanding of our world in the light of contemporary scientific knowledge. *Science, Perception and Reality* is a compilation of recent papers and articles by Prof. Sellars (who has just moved from Yale University to Pittsburgh) covering a range of topics. In one of these, 'Empiricism and the Philosophy of Mind', he is specially concerned to attack what he calls 'the myth of the given' which has, he believes, bedevilled much of traditional Empiricism. Philosophy, like Science, he argues, is a self-supporting structure which requires no indubitable foundations. *Philosophy and Scientific Realism* by Prof. Smart of Adelaide University is partly a compilation, but manages nevertheless to make a coherent unity.

The community of outlook between these two philosophers is best shown in their respective opening chapters dealing with the function of philosophy. When it comes, however, to finding a place for man in their scientific-realist world view their viewpoints diverge. Thus, Sellars sees the problem as, in the first instance, a trilemma whose three horns are '(*a*) a dualism in which men as scientific objects are contrasted with the "minds" which are the source and principle of their existence as persons; (*b*) abandoning the reality of persons as well as manifest physical objects in favour of the exclusive reality of scientific objects; (*c*) returning once and for all to the thesis of the merely "calculational" or "auxiliary" status of theoretical frameworks and to the affirmation of the primacy of the manifest image' (p. 39). This is well said. But whereas this reviewer would accept these disjunctions as exhaustive and settle for option (*a*) as being the least unsatisfactory, Sellars believes he has found an escape by invoking what Kant called 'The Kingdom of Ends'. Thus if we think of a 'person' as essentially a member of a community and so bound to oneself 'in a network of rights and duties...' the irreducibility of the personal is guaranteed, as it were, by 'the irreducibility of the "ought" to the "is"'.

Smart, on the contrary, commits himself unreservedly to option (*b*). 'My thesis', he states in chapter VI, 'is that man is a physical mechanism.' Smart, in fact, in company with some youthful followers in Australia, now represents the extreme wing of contemporary materialism. His avowed aim in writing this book is to purge us of what he calls our 'anthropocentricities'. 'Kant's so-called Copernican Revolution' he declares with, I think, some justification 'was really just an anti-Copernican counter-revolution'. If we are to understand the world aright, that is in a truly objective way, we must start by disabusing ourselves of every claim to a unique status. Thus, although the human nervous system may be more complex than other natural mechanisms we are not entitled to postulate any 'emergent' qualities as supervening at any point of the evolutionary process that connects us to the world of inorganic matter. In particular consciousness must be eliminated. To his credit, however, Smart refuses to avail himself for this purpose of the facile behaviourism of Ryle and Co. but relies instead on the 'Brain Process Theory' where he acknowledges his debt to U. T. Place (this *Journal*, **37**, 1956, 44–50). The gist of this intriguing theory is that since introspective reports are 'topic neutral' (i.e. when a subject says he is in pain we cannot tell from his protocol alone whether he is referring to something psychical or to something physical) we are justified, in the light of the neurophysiological evidence, in *identifying* the referent of such statements with something *physical* (i.e. with brain processes). For the elimination of colours and other secondary qualities, however, Smart falls back on the well tried (though to me unconvincing) arguments of operationism. More original are his attempts to get rid of anthropocentricity from our conception of time. Terms like 'past', 'present' and 'future' have to be replaced by terms like 'earlier', 'contemporaneous' and 'later'. He also speculates about the properties that would be needed for a truly cosmic language in which we could communicate with other rational but non-human species in space.

Both these books will repay study but Prof. Sellars's is strictly technical philosophy written for the benefit of his fellow professionals and his labyrinthine arguments and stodgy prose make few concessions to the layman. Professor Smart, on the other hand, is nearly always eminently readable, and his clear-cut provocative ideas should make a powerful impact on the contemporary philosophical scene.

JOHN BELOFF

The Scientific Evolution of Psychology, vol. I. By J. R. KANTOR. Granville, Ohio: The Principia Press. 1963. Pp. xxii + 387. $10.00.

Prof. Kantor's aim is to trace the development of psychology from the days of the early Ionian philosophers to the close of the scholastic period. 'Previous historians of psychology', he complains, 'immured in a dualistic culture, have imposed a spiritistic patina on what the original writers discussed.' 'The evolution of psychology', as he sees it, displays two broad 'cycles or waves'. On its first emergence it appeared as essentially a 'naturalistic science'—a mode of approach which reached its peak with Plato and his immediate followers. With Philo and Plotinus it sank into 'transcendentalism', and dropped to its lowest depth with Augustine and Aquinas. It rose once again with Hobbes and Descartes, and ultimately, owing to the influence of Pavlov, 'changed its entire character'. 'The question was raised as to why psychologists should concern themselves with consciousness and similar transcendent matters.' The outcome was behaviourism. Yet even current behaviourism, in Prof. Kantor's view, is not wholly satisfactory; in its place we must substitute what he describes as 'interbehaviourism'—a final development which will form the concluding theme of his second volume.

Some of Dr Kantor's interpretations of the leading Greek and Latin authors whom he discusses, as well as his slips in the spelling of Greek words, will set most classical historians wondering whether, instead of reading what they wrote in the original, he has not read into them what he would like them to have said. When Socrates talked about the 'soul', we are told, he was merely using the term as we should use it in speaking of 'the soul of kindness'; and Plato, so far from holding any 'theory of transcendental entities beyond the realm of space and time' (as Russell's *History of Western Philosophy* declares), was actually a down-to-earth, realistic, rational, and mathematical thinker. The prevailing notion that Plato and Socrates believed in the immortality of an immaterial soul or a world of eternal ideas, such as goodness and the other values, is the result partly of the 'inverted' explanations of post-Platonic writers like Plotinus, Augustine and Aquinas, and partly of the 'transducing translations' of nineteenth-century scholars, steeped in 'religious superstitions derived from the Judaeo-Christian tradition of their age'. No one who has read the *Phaedo* or the *Theaetetus* in the original Greek could, I think, possibly accept these revolutionary interpretations. Contrary to what Prof. Kantor supposes the belief in a life after death was not part of the Jewish tradition; and the articles of Christian creed maintain, not the immortality of the soul, but the resurrection of the body. Indeed, Aquinas and other theologians actually condemned the Platonic doctrines as heretical.

Nevertheless, there is much that is both instructive and provocative in Dr Kantor's discussion. Like Russell, he seeks to relate the outlook of the various philosophers with whom he deals to their general cultural background, and particularly the state of knowledge in other fields of science. This perhaps is the most illuminating part of his history. But his volume can hardly be recommended as a text-book for psychological students who do not already possess some first-hand acquaintance with the views of the various writers whom he seeks to fit into his highly simplified evolutionary scheme.

CYRIL BURT

Imagination. By E. J. FURLONG. London: Allen and Unwin. 1961. Pp. xiv+110. 18*s*.

Within the brief space of 100 odd pages Prof. Furlong has undertaken three major tasks: (1) philosophical elucidation of the meaning of 'imagination', (2) exposition and criticism of current theories of dreaming and artistic creativity, and (3) historical exegesis and re-interpretation of the role assigned to imagination in the theories of Hume and Kant.

The elucidatory section contains three distinctions signified by the terms 'in imagination', 'with imagination' and 'supposal'. While the first two present little difficulty, the notion of 'supposal', which plays a large part in the section on dreams, is subjected to rather puzzling treatment. Two senses of the word are distinguished and illustrated by the sentences: 'Peter, suppose (imagine) you are on the top of Everest', and 'Peter supposes (imagines) that he is really driving a space machine'. Supposal in the first case is defined as 'to mentally consider' and named 'supposal plain' to differentiate it from the second case in which supposal is equated with 'falsely think' or 'falsely believe'.

Apart from a minor confusion which arises when the author subsequently categorizes dreams as instances of 'false supposal', a serious difficulty is encountered with the ambiguity of such phrases as 'falsely think' and 'falsely believe'. I may, for example, truly believe that Michigan is in Canada. Do I 'falsely believe' because the proposition is itself false?

The distinction between 'supposal plain' and supposal as 'false belief' could presumably be made by talking of imagination with and without belief in the reality of what we imagine. Indeed, we are told that supposal plain passes into false supposal '...when Peter forgets that his expectations are copies: he takes the copy for the real'. Illusions and hallucinations '(in so far as they deceive)' are also given as instances of false supposal. It seems that the phrase 'false supposal' is to serve the purpose of characterizing a mental state. Can it at the same time connote the truth value of the proposition supposed or believed?

When the author describes dreams as instances of false supposal the phrase is presumably intended to mark a particular kind of mental state. Later in the text, however, we are presented with a hypothetical dream '...that Australia has won the test match. We wake up, turn on the wireless, and the announcer confirms our dream.... Hence, in this dream we could not be said to have falsely supposed.'

The conclusion that there is a confusion here brought about by the ambiguity of such phrases as 'false belief' is hard to avoid, and the position remains unclear throughout the text.

In the section on dreams some cogent common-sense arguments are advanced against the kind of dream theory advocated by Malcolm. The notion of a frontier between sleep and wakefulness with no intermediate state is adequately disposed of. A chapter on images contains speculation about empirical matters with no reference to the wealth of actual data accumulated by psychologists. 'It may be useful' we are told 'to take account of the experience of a blind man—blind from birth. How far would he be inclined to use the image terminology?' The author invents his blind man rather than finding one so afflicted and asking him.

Two small concluding chapters on Hume and Kant contain a clear exposition of their respective viewpoints. It is suggested that Kant's 'productive imagination' may be regarded as 'supposal plain' but the matter is not pursued. The reader is left with the impression that far too big a task has been attempted in far too small a book. The style is urbane but the content lacks the subtlety and rigour that one has come to associate with the best of modern philosophy.

J. K. CLARKSON

Affect, Imagery, Consciousness. Volume II. *The Negative Affects.* By SILVAN S. TOMKINS. New York: Springer; London: Tavistock. 1963. Pp. 580. 65*s*.

In the first volume of his treatise, the author propounded his theory of the affect system as a primary source of human motivation, and as distinct from and independent of the drive system. Although the latter signals information which excites narrowly directed types of behaviour such as are necessary for self-maintenance, the former is responsible for more general and enduring kinds of activity, rewarding them through the positive affects and punishing them through the negative. The nature of the affect experienced was linked to increase or decrease of stimulation. Sudden increase was hypothesized to produce startle or fear; less sudden increase, excitement or interest. Prolonged high-level stimulation resulted in distress, more sustained but less toxic than fear. Sudden decrease was said to give rise to laughter, and less sudden decrease, to joy. But the nature of the affects was also influenced by the removal, reduction or interruption of existing affective states.

In his first volume the author also expounded the nature and sources of the positive affects. In this volume he has turned to the negative affects: first, distress–anguish; secondly, shame–humiliation, with which are associated shyness and guilt, all being in opposition to interest and joy; and thirdly, contempt–disgust, the latter being closely associated with the primitive drive of food rejection. The author attaches great importance to humiliation and its avoidance in the development of personality, linking excessive humiliation with the paranoid personality.

There is much here to interest those who wish to consider carefully the very detailed discussion of these complex emotional states in all their intricacies, their beginnings in the infant and their further development in the child in reaction to parental treatment and other forms of environmental situation. But the presentation of the argument appears unsystematic and is often difficult to follow; it is hard to grasp what the author intends. Moreover, the general theory of affect seems to be based largely on *a priori* reasoning, too often illustrated only by casual observations or by lengthy citations of, for instance, the history of the 'evil eye', the writings of Dostoevsky and Strindberg. In places case histories and experimental data are included, though the value of the latter is decreased by the absence of references (presumably to be quoted in the final volume of the publication). But again these appear to be illustrative, rather than constituting the basis of hypotheses of the nature and development of the emotional states. Of course there is a shortage of reliable experimental data, it being difficult to design satisfactory techniques of observation and assessment. Nevertheless, it is a pity that so much of the analysis is presented as if its statements were in fact proved, rather than in the form of hypotheses which could be put to some experimental test.

The reader who is prepared to devote much time and effort in unravelling the author's argument, and in extracting that which is valid from that which is pure assumption, will doubtless profit from so doing. But it is not a task to be lightly attempted.

M. D. VERNON

Current Perspectives in Social Psychology; Readings with Commentary. Edited by E. P. HOLLANDER and RAYMOND G. HUNT. New York and London: Oxford University Press. 1963. Pp. xi + 557. 36*s*.

When we have read the various editions of Newcomb and Hartley's *Readings in Social Psychology* we know something of what social psychologists have found out about life. We know what people actually did in the 'extreme' situation of a concentration camp, or what the influence of sex and status are in the deliberations of juries. Hollander and Hunt in this volume hardly touch this sort of matter but, as their title suggests, focus on second-order questions of method, conceptual framework, and theory. In their own words, they aim to give attention to the broadest range of prominent and productive viewpoints which characterize the field today.

The actual readings consist of fifty papers selected from publications that appeared between 1950 and 1961. Fifteen contributions are from the *Psychological Review*, twelve from other psychological journals, eleven from sociological and anthropological journals, eleven from books, and one appears to be original. (There is a lone European voice, Hilde Himmelweit's. It is a pity that she has been represented by an abridgement from a UNESCO publication, which is over-condensed and anyway only an interim account of data.)

The papers are organized in eight sections under currently seminal headings, viz.: Culture, Learning and Group Identification; Personality and Society; Language and Communication; Person Perception, Interaction and Role; Attitudes and Cognition; Normative Behaviour, Conformity and Intra-Group Processes; Leadership, Power and Innovation.

Each section is introduced by a commentary by the editors. These are more than a justification for the authors' names on the cover. They give an over-view of thinking in each area in general, and show how it is represented by the present papers. They provide extensive cross-references to other papers in other sections of the volume, not only in terms of content but with respect to their intellectual history. For example, they point out that Bruner's *On Perceptual Readiness* contribution was the stimulus for Berkowitz on *Judgmental Processes in Personality*; Brown on *How Shall a Thing be Called?*; Cantril on *Perception and Interpersonal Relations*; and Steiner on *Accuracy of Social Perception*. Further references to outside books and articles are freely made.

It is seen then that this book in no sense replaces that of Newcomb and Hartley, but is the next step for the student in this field. It documents the next step that social psychology itself has taken.

It is a paper-back. It is, however, in medium octavo format, excellently bound and printed. At the price it is a good 'buy'.

HALLA BELOFF

Personality and Social Systems. Edited by NEIL J. SMELSER and WILLIAM T. SMELSER. New York and London: J. Wiley. 1963. Pp. xi+660. 68*s*.

Although one is sometimes tempted to think that we may be getting too many compilations of previously published work by different hands, it would be quite unjust to regard the present volume as excess luggage. It should be, on the contrary, an extremely useful book. Sixty articles are included, and these cover a very wide range and are selected from many different journals. There are six main sections in the book, sections two to six being further subdivided. The first section consists of a fairly short introductory chapter by the editors, in which they outline their view of the field which the selections cover; and each of the other five sections has a short preliminary outline. Section two deals with theoretical and methodological issues. The reviewer found this section perhaps the most interesting, since it is in general concerned with the kind of problems which arise when psychology and sociology come into contact. The article by Reinhard Bendix on compliant behaviour and personality, reprinted from the *American Journal of Sociology*, 1952, raises basic questions which cannot fail to be of concern to the psychologist and sociologist interested in relationships between personality and culture.

The third section of the book deals with personality variables as dependent in relation to independent social system variables. Among the more familiar pieces here is R. K. Merton's 'Bureaucratic structure and personality'. Among those which were new to the reviewer were one by George de Vos on guilt towards parents, achievement and arranged marriage among the Japanese, reprinted from *Psychiatry*, 1960, and one by Lee and Schneider on hypertension and arteriosclerosis among executives, reprinted from the *Journal of the American Medical Association*, 1958. If the findings of the latter article are to be depended upon, they suggest that we may have been too ready to assume a tendency for the life of the business executive to dispose men to these disorders. The comparison was, admittedly, made between executives and non-executive white-collar workers, but the findings reported at least enjoin caution.

The fourth section stands the previous section on its head, so to speak, and deals with social system variables as dependent variables and personality variables as independent variables.

Perhaps the most interesting subsection here is that concerned with personality and social change. One short but interesting article which the reviewer for one would have been most unlikely to come across elsewhere is by E. M. Bruner on 'Cultural transmission and cultural change', reprinted from the *Southwestern Journal of Anthropology*, 1956. The fifth section deals with the 'combination of personality and social variables to account for empirical regularities', while the sixth section considers the relation of social structures to the 'rehabilitation' of personality—for example, the article by Doris Gilbert and D. J. Levinson on 'Ideology, personality and institutional policy in the mental hospital'; and Sol Ginsburg's article on social workers and social therapy, reprinted from *Social Casework*, 1951. The reviewer feels even less qualified to comment on these last two sections than on the earlier ones.

To repeat, this is an exceptionally interesting and varied selection which should be an invaluable source of food for thought for students, teachers and researchers in psychology and sociology. There are just two things to which one might take exception. The fact that the pages are double-columned will not please those who dislike this form of printing; and references are given in different ways, according to the system used by the journal concerned. Consistency would have been an advantage, though perhaps the time and cost involved might not have been really justified.

D. GRAHAM

Personality Development and Psychopathology. A Dynamic Approach. By NORMAN CAMERON. Boston: Houghton and Mifflin. 1963. Pp. xxi + 793. $8.75.

Although this is a very long book, and occasionally repetitious, Norman Cameron has never been an 'over-includer'. He proceeds with admirable clarity on the systematic course indicated by the title. Only rarely, when he is leaning particularly heavily on psychoanalytic doctrine rather than on his vast and perceptive clinical experience, does his customary good sense seem less in evidence.

'The mature super-ego...amounts to the presence of society's moral representative, not only right in the home, but right in one's most private, personal experience. This presence, at its best, guides a person's life along socially approved lines.' Surely society generally approves the modal patterns of conduct and pays scant regard to individual differences, often with disastrous results?

The account of personality development is a fairly orthodox psychoanalytic one, including what is always alleged to be the convenient shorthand of such terms as super-ego, anal fixation and the oedipal child, which the reviewer always feels constrained to translate freely into inconvenient long-hand. The author's account is, indeed, so good that one hopes that it will not be necessary for it to be done again without at least some tentative indications as to how the numerous suppositions might be confirmed or refuted. We are told, for example, that depressive personalities show 'an unusual devotion to eating and drinking for their own sake, addiction to smoking and chewing...'. Somebody could find out whether or not this is true.

The egocentricity of psychoanalytic theory is less in evidence here than is often the case; but it is not absent. Other people tend to be seen as need-satisfiers. 'Needs for human companionship, for freedom of action...play significant roles in preserving personal integrity.... Why do we leave them out? We confine ourselves to sexual and aggressive drives because under ordinary conditions these are the drives which create major problems in behavior pathology.' Two pages later (p. 134): 'Direct sexual and aggressive drive satisfaction requires the reciprocal behavior of another person if it is to be complete. This is not true of satisfying hunger or thirst.' It is psychological isolation from other persons as persons, not from human bodies in bed or in a boxing ring, that contributes to the determination to commit suicide. The need for human companionship is usually left out, not because it is unimportant, but because we do not know how to talk about it.

'Recreation, creative activity, and sex intercourse' are spoken of as 'temporary regressive episodes' as though that was all that they were. Creative activity in particular may contain a temporary regressive element; but it may, in the main, be progressive and integrative at a higher level than had previously been attained.

These rather general cavils, to which a few minor ones could have been added, are far outweighed by the numerous excellent features, and are by no means representative of the author's views. The accounts of symptom formation and of the predominance and operation of different defence mechanisms in different clinical syndromes are especially clear and helpful. Altogether one feels unusually confident that psychoanalytic insights are being applied by someone who really knows psychiatry in all its aspects. Accordingly this book can be recommended to all those training to become psychiatrists or clinical psychologists as an authoritative treatment of its subject matter by one of the most versatile and unbiased psychiatrists of the present time.

G. A. FOULDS

Modern Perspectives in Child Development. Edited by A. J. SOLNIT and S. A. PROVENCE. New York: International Universities Press; London: Bailey Bros. and Swinfen. 1964. Pp. xxi + 666. 112*s*.

This volume, produced to honour Milton Senn on his sixtieth birthday by his present and former colleagues and students, demonstrates by its breadth of coverage the scope of his concern in the fields of child development, paediatrics, child guidance, medical and general education; while the nature of his influence can be seen in the close association—amounting at times to fusion—of medical and psychiatric points of view.

Of the forty-two contributors, about one half are paediatricians. It is interesting to see how this profession, so influential in America, is examining its own assumptions and methods, taking increasing cognisance of children's personal needs, and seeking to define its role in relation to psychiatry.

Most of the remaining contributors are psychiatrists or psychoanalysts; there are two or three representatives each of psychology, education and social work. The reason for the comparative neglect of the psychological contribution to knowledge of child development appears to be that the predominantly clinical team is less interested in the quantitative approach, which of necessity confines itself to the measurement of variables abstracted from the whole, than in the description of the total personality, or at least of whole areas of functioning which largely defy the psychologist's present tools.

The most substantial chapter is a 100-page essay (plus 25 pages of bibliography) by Dane Prugh discussing children's illnesses in the light of psychosomatic concepts. Most of the papers are less ambitious, but by no means inconsiderable. Of particular interest to psychologists are one by Anna Freud on the role of regression in mental development; one by John Doris on the evaluation of the intellect of the brain-damaged child; a review by Richmond and Caldwell of present knowledge concerning the effects of child-rearing practices; two chapters on child guidance, and several on the attitudes of physicians to such questions as school avoidance, adolescents' problems and the hospitalization of young children. There is also a case study, illustrating the individual-centred approach of the Yale Longitudinal Child Study initiated by Kris and Senn, which might have been made to yield more hypotheses than are produced to demonstrate its fruitfulness.

Most stimulating of all is an essay in which Lawrence K. Frank envisages the science of human development as an emergent discipline, destined to transcend the barriers between the contributory sciences and to replace their borrowed conception of linear causal sequences by new concepts of interrelationships between the many simultaneous ongoing processes that constitute a living organism. This chapter and the book as a whole present many challenges to our current assumptions.

TERENCE MOORE

The Family and Human Adaptation. By THEODORE LIDZ. London: The Hogarth Press and the Institute of Psychoanalysis. 1964. Pp. 120. 25*s*.

This book consists of three lectures: the Mona Bronfman Sheckman Lectures in Social Psychiatry at Tulane University in 1961. The lecturer addressed himself to psychiatrists, anthropologists, psychologists, sociologists and members of allied professions. The book would provide thoughtful reading for a train journey, but, although full of interesting and interestingly presented ideas, is somewhat disappointing. The reviewer's impression is that professionally trained readers of the kinds envisaged by Dr Lidz would read it with enjoyment and interest but without finding anything they did not already know, or having their outlook changed in any important way. However, beginners and students in the subjects mentioned would find the book a useful formulation, in a short space, of numerous ideas about the family and its functions, problems and changes, which would be the starting-points of various associated trains of thought and the pegs on which to hang many of the concepts necessary for their work.

Perhaps this is the way in which the book should be viewed, because it is simply a group of lectures bringing a certain outlook into focus. The first lecture is about the family and human adaptation in the scientific era, the second about family organization and personality structure, and the third about the family, language and ego functions.

In reading the third lecture one does not want to attribute to the author the intention to claim that schizophrenia is produced by conditions in the family which led to failures in learning clear and socially adapted forms of verbalization, conceptual thinking and problem solving, but what he says often comes near to giving this impression. Surely unless the psychodynamics of the problems of schizophrenia can be focused clearly we are simply talking like the 'conscious' psychologists who deal with individual development as though it had no unconscious dynamics, but was based on the processes of conditioning.

Schizophrenia is, according to Dr Lidz, 'a gross failure to attain or maintain ego integration', and he adds that 'schizophrenic patients received a faulty and confused grounding in linguistic meanings, as well as in other instrumental techniques, that both limited their adaptive capacities and permitted them to escape from insoluble conflict or irreconcilable contradictions by abandoning the meaning system of their culture' (p. 5). This leaves the fundamental psychodynamic problem open. The illogicalities of the schizophrenic could not be de-conditioned without a transformation of his unconscious psychodynamics.

R. W. PICKFORD

Leonardo da Vinci; Psychoanalytic Notes on the Enigma. By K. R. EISSLER. London: Hogarth Press. 1962. Pp. xx + 375.

This book falls into two main parts. The first consists of a somewhat polemical discussion of the relevance, in principle, of a psychoanalytic approach to art, irrespective of actual mistakes that have been found in Freud's discussion of Leonardo. In the second half, Dr Eissler embarks on a lengthy consideration of the personal and sexual aspects of Leonardo's life in relation to his artistic output.

The book is likely to appeal only to those with highly specialized interest in the subject-matter. For Dr Eissler's material is not well marshalled and his technical vocabulary and some lengthy sentences make the going heavy.

There are incidental rewards. For example, when referring in passing to Goya, Dr Eissler remarks on his capacity to paint unconscious aspects of his sitters' personalities. Does this bear on the rather strange fact that Goya was a fashionable painter in spite of the traits of cruelty, vanity, mean spiritedness and the like that stand revealed in some of his portraits? Did his eminent sitters not see the evidence of these traits?

JOAN WYNN REEVES

The Evolution of Psychiatry in Scotland. By SIR DAVID HENDERSON. Edinburgh: Livingstone. 1964. Pp. vii + 300. 32*s.* 6*d.*

This delightful book is written by the most eminent of living British psychiatrists, who has become a part of the history of his speciality in his own lifetime. The first half of the book contains biographies of leading reformers and psychiatrists and an account of the development of the famous Scottish Royal Mental Hospitals. The second half is mainly autobiographical and ends with the author's views on the future of psychiatry.

It is interesting that the book was published almost at the same time as the 1963 Harveian Oration by Sir Aubrey Lewis. The difference of approach and emphasis between Henderson and Lewis is almost identical with that between their predecessors Clouston and Maudsley nearly 100 years ago. Henderson is enthusiastic about the wide application of psychiatric knowledge to society and boldly claims that crime is a disease, while Lewis is cautious and wishes to confine psychiatry to its many unsolved clinical problems.

Whether one agrees with Sir David's enthusiastic optimism or not, this book can be unreservedly recommended to all psychiatrists, psychologists and others who are interested in the state of British psychiatry, past and present.

FRANK FISH

Psychiatry for Students. By DAVID STAFFORD-CLARK. London: Allen and Unwin. 1964. Pp. 277. 35*s.*

Dr Stafford-Clark's Pelican on '*Psychiatry Today*' was so admirable and vivid that teachers of psychiatry to medical students have been waiting for the text-book they hoped he would write: it turns out to be a qualified success. However, even the apparent shortcomings may be intentional. The student is provided with strikingly little information, which means that glaring omissions and a high degree of selectivity occur. This economy in factual content may be based on experience that students have a limited tolerance for new information. The psychodynamic hypothesis is presented in a chapter just four pages long, on the basis of two concepts, repression and the unconscious mind.

The author holds that the chief contribution which psychiatry can make in training of future doctors is the psychosomatic orientation, and he urges on his readers the view that psychological illness is 'fundamentally an aspect of all illness'. Aubrey Lewis and others have warned psychiatrists against making the claim that more than other physicians they have a premium on humanitarian feeling; in claiming the psychosomatic approach as the psychiatrist's main strength in the medical curriculum, the author disregards the likelihood that the general physician will resent the usurpation, and question the accuracy of the assertion that discovery of the effects of excessive and undischarged emotion has thrown light on the aetiology of illness in general. The student reader is told in the first pages that the author's intention is to concentrate on those aspects of psychiatry which are of relevance to any practising doctor. Thirty-seven brief case histories are presented as illustrations. The induction of hypnosis and indications for its use are set out in eleven pages.

An individual system of classification is used. Functional neuroses comprise hysterical and obsessive-compulsive reactions. Anxiety states, depressive illness and mania are grouped as affective disorders. Schizophrenia is in a category by itself. Acute confusional states are grouped with 'other psychiatric emergencies'. Puerperal psychoses form a separate category. The personality disorders are classified under the general heading 'disorders due to failure of normal development', with the amentias and the instinctual anomalies (the latter including alcohol addiction).

Another unusual feature is the use of exhortation. Readers are enjoined to be compassionate to patients, and are told that psychiatry contributes to medicine the underlining of a single fundamental truth, the essential wholeness and dignity of man. Such frank attitude-moulding is rare in medical text-books, but may well be effective, and the author clearly holds that the teacher should attempt to modify students' values in a direction conducive to good doctoring.

H. J. WALTON

Hypnosis and Suggestion in Psychotherapy. By H. BERNHEIM. Translated from the second revised edition by CHRISTIAN A. HERTER. New York: University Books. 1963. Pp. xvi + 428. $10.00.

Clinical Correlations of Experimental Hypnosis. Edited by M. V. KLINE. Springfield, Ill.: Charles C. Thomas. 1963. Pp. xv + 524. $15.75.

The first of these two handsomely bound books represents a landmark in the emancipation of psychiatry and psychology. Appearing in 1886, it was translated into German (by Sigmund Freud) and into English in 1888 under the title, 'Suggestive Therapeutics'. Because, say the publishers, it has been out of print for 75 years (actually it was reprinted in 1947 by the London Book Company), they have republished it under a new title. The book contains a long series of case reports of patients whose hysterical complaints were terminated by suggestion. The controversy between Charcot and Bernheim over the nature of hypnotism, and the vindication through time of Bernheim's insistence that all the phenomena of hypnotism arise through the operation of suggestion, form part of psychology's history. That suggestion is the leading feature common to all the warring schools and exponents of psychotherapy will not be lost upon the reader as he thumbs through Bernheim's case reports under a new title. Here is an opportunity for new departments to acquire an historical classic.

The second volume illustrates the twentieth-century way of doing things. The experienced editor compiles a collection of specially written essays together with research articles previously published in journals as diverse as the *Lancet* and the *Journal of Experimental Psychology*. The range of quality, of sophistication and naïvety, and of topics covered, are equally diverse. The phenomena of hypnotism offer an enduring challenge to research. Anyone wishing to gain acquaintance with the present state of knowledge could certainly do worse than to read this book.

IAN OSWALD

Developmental Dyslexia. By MACDONALD CRITCHLEY. London: Heinemann. 1964. Pp. xi + 104. 25*s*.

Speech and Language Delay; a Home Training Program. By R. RAY BATTIN and C. OLAF HAUG. Springfield, Ill.: Thomas. 1964. Pp. xi + 71. $4.50.

This monograph is an expansion of Dr Critchley's celebrated Doyne Memorial Lecture, *Inborn Reading Disorders of Central Origin* (1961). In it he describes the development of thought about developmental dyslexia and reviews a vast amount of literature spanning approximately four centuries. Inevitably, the bias of his interest is manifest in the way in which different aspects of developmental dyslexia are treated. The paediatricians will probably feel that there is insufficient discussion of the clinical problems presented by developmental dyslexia; psychologists are likely to think that his account of tests of reading and writing ability and achievement are inadequate and psychiatrists may find his chapter on 'Psychiatric repercussions' too limited.

It is to be remembered, however, that the monograph does not set out to be a practical guide to the diagnosis, assessment, management and remedial teaching of children suffering from developmental dyslexia but seeks 'to trace the growth of our knowledge of this condition and to describe the conflicting ideas as to nature and causation'. The author has succeeded in doing just this in a masterly way within eighty-nine pages to which are added another twelve and a half pages of bibliography. The book is beautifully produced and well illustrated. It will be an invaluable reference book for a long time.

Speech and Language Delay is a dogmatic little book designed for American parents of children with slow speech development. It may have a limited place as a source of advice to parents when their children require treatment by a speech therapist, but it should be recommended only by those who have read it carefully, and then with care and discretion.

T. T. S. INGRAM

Color; a Guide to Basic Facts and Concepts. By ROBERT W. BURNHAM, RANDALL M. HANES and C. JAMES BARTLESON. New York and London: Wiley. 1963. Pp. xi + 249. 70*s*.

This book is a report of the Inter-Society Color Council Subcommittee for their *Problem* 20: *Basic Elements of Color Education*. It was distributed to members of the I.S.C.C. free, but is also on sale from the publishers.

It may be useful for readers of reviews to note the aims embraced in the plan evolved by this Subcommittee. They were as follows:

(1) To prepare a detailed outline of the facts of colour in which each statement could stand by itself out of its context.

(2) To code the outline in such a way that ready reference could be made to any part of it.

(3) To document the outline adequately.

(4) To cite primary sources whenever possible.

(5) To adopt a set of colour terms near to technical and commercial usage but also carefully and consistently defined.

(6) To insert an index which refers to all colour terms and concepts in the outline.

(7) To include many pictorial illustrations, in colour whenever possible (and indeed these are splendid).

In addition, the outline was planned to serve as a course for instruction in the study of colour, as a compact and reliable reference book for a student who has completed such a course, and as a basis for writing textbooks on colour.

These demands resulted in the production of a very interesting and impressive, but somewhat unusual book. We do not need any page number to refer to something in it, but only the appropriate number of the statement in question, e.g.: '2.1.3 All sources of light may produce color stimuli', and so on. Page references, however, would be needed in order to find out the number of a statement required. The reduction of the whole subject of colour in all its aspects and ramifications to a long series of interrelated assertions, each designed to be a true statement if taken out of its setting, makes for rather hard reading. Nevertheless, many or most of the sentences and paragraphs are most admirably constructed, and are perfect verbalizations of the concepts and facts presented. Many of these facts and concepts are, as everybody knows who has worked at all seriously with colour, very subtle and 'tricky' to express, and the sophisticated reader will look through the book with delight at every new discovery of its many precise and lucid statements.

There is a little difficulty about many of the statements, however, and it is not easy to see how a sentence like '11.3.1c.2a) The HRRT contains 21 PIC plates of four general types', could stand by itself, out of its context, except to a person who had read, understood and remembered statements '11.3.1a.2.b).1)' and '11.3.1c.2', beforehand, because it is in them that 'PIC' and 'HRRT' are defined. It is clear, however, that to write out all these abbreviations in full would considerably increase the length of the book.

The authors will forgive the reviewer for this slight levity, because the book is a masterly piece of clear presentation and accurate statement, like the Statue of Liberty itself, or the Washington Memorial.

R. W. PICKFORD

Laboratory Studies in Operant Behavior. By J. MICHAEL. New York and London: McGraw-Hill. 1963. Pp. xxv+89. 23*s*.

This is a laboratory manual designed for undergraduates and is devoted to the study of operant behaviour in the rat. It is intended for use in close conjunction with Holland and Skinner's programmed learning text *The Analysis of Behavior*. It describes nine experiments ranging from simple magazine training to relatively complex behaviour in which, for example, the rat learns to pull down a swinging bar in order to turn on a light, only after which may it press a lever to obtain food. Clear instructions are given on the procedure to be followed in each case and space is given for the individual student to enter the results obtained. The spiral-spring binding is adequate, and the format attractive.

What present application this programme might have in this country is doubtful. Few if any departments are adequately equipped to allow each student to have his own Skinner box, even if it were desired to lay so much emphasis on operant conditioning. Moreover, there might be difficulties over the use by students of motivating techniques such as food deprivation and it might be held necessary for them to obtain Home Office licenses for the purpose. Perhaps some of the experiments described in this useful manual could be performed as classroom demonstrations but this would obviously detract from the value which the author clearly (and rightly) believes is to be derived from intimate contact of experimenter with the behaviour of the subject under test.

P. L. BROADHURST

The Design of Electric Circuits in the Behavioural Sciences. By T. N. CORNSWEET. London and New York: Wiley. 1963. Pp. xi+329. 68*s*.

If ever a book arrived at the 'right psychological moment' it is this one. For now as never before the experimental psychologist is likely to find himself using pieces of machinery which for sheer complexity may well put the brass instruments of a bygone era in the spinning jenny class.

In this 'do it yourself' guide Cornsweet shows a real understanding of the difficulties facing those who are both unsophisticated in matters electrical and short of technical aid. With some

70,000 words helped out by over 250 plates and diagrams this book, small enough to go in the pocket of a laboratory worker's overall, takes the reader from the elegant simplicity of Ohm's law to the complexities of a modern cathode-ray oscilloscope.

Its aims are threefold—to teach the simple principles of electricity and electronics, to give practical advice on the construction of electrical apparatus, and to show how particular experimental needs may be most suitably met.

It succeeds in the first two incomparably well. It is only in its third intention that the book may be faulted on presentation and content. Improvements could be made in both the depiction of circuits and range of problems dealt with. The 'disentanglement' of the circuits presupposes a level of perceptual learning in excess of that likely to have been acquired by the novice in the earlier chapters of the book. The use of such visual aids as a colour code for primary and secondary circuits might well have made the task easier. Apropos of the range of problems it is disappointing to find that 'behavioural sciences' means 'rat psychology'. The space devoted to circuits for timing rats, feeding rats, shocking rats, and counting rat responses far exceed that given over to more human affairs. Obviously much of this will be relevant for experiments on humans but while sections on, for example, such admirable topics as how to prevent faeces from short circuiting electric grids may well delight those who work with rodents, they will be cold comfort for the psychologist who still thinks in terms of high speed tachistoscopes, memory drums and kindred display gadgets, audiometry, and the recording of multiple data on magnetic tape.

But these are relatively minor criticisms. On the whole this is a unique and excellent book and one which should date rather less quickly than most textbooks for psychologists.

N. F. DIXON

Psychology's Impact on the Christian Faith. By C. EDWARD BARKER. London: Allen and Unwin. 1964. Pp. 220. 28*s*.

This book has been written by a Methodist minister who, after a period of physical illness, took a training analysis and now practises as a lay analyst.

It is not surprising, therefore, that he has nothing to say about scientific psychology. It would have been less misleading if he had called his book: 'The Impact of Psychoanalysis (Klein–Fairbairn variety) on the Christian Faith (evangelical variety)'. This particular field has received fairly extensive treatment already over the last five or ten years, and Mr Barker's book adds little to its predecessors in the field, though it may well be read with profit by evangelically minded ministers of religion. Most psychologists will be infuriated by the psychoanalytical and theological speculations which are adduced as matters of fact.

RALPH HETHERINGTON

Behaviour Theories and the Status of Psychology. By GEORGE WESTBY. Cardiff: University of Wales Press. 1963. Pp. 32. 3*s*. 6*d*.

A wise and witty inaugural lecture from the first Welsh Professor of Psychology; who warns against over-confidence; castigates the deviant heresies of ultra-empricism, ultra-operationism, inveterate reductionism, and the chaotic chorus of amateur depth psychologists, existentialists and their like; praises the modest, cool, sceptical scientists who increasingly dominate the psychological scene; provides grounds for optimism; and all in all presents his audience (and now his readers) with good value for an hour of their time or a few shillings of their cash.

L. S. HEARNSHAW

Vol. 55. Part 4 November 1964

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY

BORIS SEMEONOFF

AIDED IN THE SELECTION OF PAPERS BY

C. BURT

AND

F. C. BARTLETT
E. A. BOTT
D. E. BROADBENT
R. S. CREED
J. DREVER
G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
M. D. VERNON

ASSISTANT EDITOR

R. M. McKENZIE

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W. 1
AMERICAN BRANCH, 32 EAST 57TH STREET, NEW YORK 22, N.Y.

Price Twenty-five Shillings net (U.S.A. $4.25)

Brit. J. Psychol.

Preparation of Papers for the British Journal of Psychology

The *British Journal of Psychology* subscribes to the *Fair Copying Declaration* of the Royal Society under the terms of which reproduction of material published in this Journal is permitted.

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. All papers, including reviews, must be typed (double spacing) on one side of the paper only and the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of most articles published from Vol. **47** (1956) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

2. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs etc. is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered caption of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following example:

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.
RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* **50**, 83–96.

Note that World List abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

British Journal of Psychology, Vol. 55, No. 4

Professor C. W. VALENTINE

(*Facing p. 385*)

Brit. J. Psychol. (1964), 55, 4 *pp.* 385–390
Printed in Great Britain

CENTRAL LIBRARY UNIVERSITY

OBITUARY NOTICE

CHARLES WILFRED VALENTINE, 1879–1964

By the death of Professor Valentine on 26 May both British psychology and British education have sustained a grievous loss. In spite of failing strength he was active almost to the end. His last book was published when he was eighty-two; and I have, in his own handwriting, an article drafted only a few months before his death.

Charles Valentine was the son of the Rev. Henry Valentine, a Methodist minister. As a result of the circuit principle the family was continually on the move. Charles was born at Runcorn; he began his education at the High School in Nottingham, he continued it in Preston Grammar School, and eventually became an Exhibitioner at University College, Aberystwyth, where he took a pass degree. Later he entered Downing College, Cambridge. Here he switched from the physical sciences to the Moral Sciences Tripos, and graduated with double first-class honours in Philosophy and in Psychology. On Ward's advice he went to Germany with a view to studying the new experimental methods of Wundt and his disciples; and it was as a fellow-student in Külpe's laboratory at Würzburg that I first met him. On his return Myers offered him as a post as Demonstrator at Cambridge; but in the end he decided to accept a lectureship at St Andrews, where he gained his doctorate. In 1914 he was appointed to the Chair of Education at Belfast, and five years later he became Professor of Education in the University of Birmingham—a post which he occupied for twenty-seven years.

Few living psychologists have ranged over so wide a variety of subjects, or worked with so many different techniques. Besides his numerous contributions to child study and education, he has published original investigations on optical illusions, the genesis of laughter, the appreciation of painting and music, and the religious philosophy of Lotze. According to the varying demands of his problem, he was ready to use either the old-fashioned methods of introspection or the latest techniques of the experimental laboratory. He was a shrewd observer, a skilful interviewer, particularly ingenious in the construction of questionnaires, rating scales, and standardized tests, and one of the first to adopt what were then the novel devices of statistics.

Of his experimental researches the earliest was an intensive study of the familiar 'horizontal-vertical illusion'. The investigations were carried out first in the Cambridge laboratory, and then with students and school children at St Andrews and Dundee. The main object was to check Lipps's explanation in terms of *Einfühlung*; and the results, based on over 20,000 judgements, showed clearly that Lipps's hypothesis was quite inadequate to account for the facts.

This research was followed by others dealing more directly with visual and auditory appreciation in the field of aesthetics. The first was a series of experiments on preferences for chords and discords. Once again a carefully designed experimental technique was combined with detailed introspective records. Both children and adults served as subjects; and the latter included Prof. A. E. Taylor, whose introspections proved particularly instructive and served as a starting-point for later researches. The fourfold classification of types—'objective', 'subjective', 'associative', and 'character' types—already noted by Bullough in his work on the appreciation of colours, was found to apply, with minor modifications, to the appreciation of musical sounds; and the general results appeared very largely to confirm the theory put forward by Myers, namely that 'the preferences evinced by individuals depend largely on frequency of association, or, in one word, familiarity'. Similar experiments were carried out on the appreciation of form, colour, and pictures; and a year later, at the suggestion of Prof. Stout, Valentine brought together his chief conclusions in his first book, *An Introduction to the Experimental Psychology of Beauty* (1913), which in those days cost only sixpence.

In education his first research consisted in a study of the relative merits of the phonic and the look-and-say methods in the teaching of reading; for this purpose he used the novel device of teaching children to read English prose transcribed in Greek letters. Shortly afterwards he published an *Introduction to Experimental Psychology in Relation to Education*. This was essentially a hand-book of practical exercises, and appeared at a time when few educationists had as yet realized the value of the new experimental approach in the training of teachers. His object was not merely to illustrate the uses of scientific techniques in educational research, but also to show how a psychological outlook could help the ordinary teacher in the everyday work of the classroom. From time to time the book was revised and amplified; and its lasting merits are demonstrated by the call for a fifth edition only a few years ago.

The New Psychology of the Unconscious (1921) dealt with an entirely different topic The book was written soon after the first World War in the days when Freud and his theories were still a novelty. Valentine's exposition sought to re-interpret the tenets of psychoanalysis in terms of orthodox psychology, and to bring its essential elements into line with the views of contemporary British writers like Stout, McDougall, Rivers and Myers. He was, I believe, the first to draw attention to the fact that some of the basic points of Freud's theory of dreams and censorship are to be discovered in Plato's discussion of dreams in *The Republic*.

His book on *The Reliability of Examinations* (1932) embodied the results of an extensive investigation into the examinations for entrance to secondary schools and universities. The book had considerable influence in official as well as academic circles, and was even quoted in the House of Commons by the Parliamentary Secretary of the Board of Education. The widespread interest which now centres on the examinations at 11-plus seems to have diverted attention from the equally important chapters on university scholarships and exhibitions. His demonstrations of the low reliability of the judgements on which such awards were based, particularly at the universities, came in those days as a starting revelation. Even at Oxford and Cambridge 20% of the scholarships and exhibitions, we are told, were 'misplaced' 'they might better have gone to the commoners who got firsts instead of to the exhibitioners or scholars who got seconds or something lower'. A few of his recommendations have since been accepted; but he was himself continually urging the need for a new inquiry dealing with post-war conditions.

During the war he was largely occupied in preparing memoranda for the War Office on methods of selecting and allocating recruits, on problems of military training and instruction, and of the maintenance of morale. A vast amount of data was collected and analysed; and the main results were subsequently published in book form under the title of *The Human Factor in the Army* (1943).

Prof. Valentine's *magnum opus*, however, is unquestionably his large volume on *The Psychology of Early Childhood.* This was completed before the war began, but did not appear until 1942. His aim was to trace the main aspects of the mental development of the child during the entire pre-school period, that is, from birth to the age of 4 or 5. The work is based primarily on the author's daily observations on his own five children, continued over a period of nearly twenty years and supplemented by information on certain special points obtained from former students and other observers. The material thus collected is analysed and compared with records for individuals or for groups published by previous writers from Darwin and Preyer to Stern, Gesell and Blatz. At the time he wrote, the mass of facts embodied in these various publications had, as he says, never been systematically collated or organized into a single synoptic review. This was followed by several smaller books on child psychology, including manuals of intelligence tests for children aged 2 to 8 and (a little later) for those aged 8 to 15, and a set of *Reasoning Tests* for adults up to university level. But of all his numerous publications that which has enjoyed the widest circulation is the comprehensive volume entitled *Psychology and its Bearing on Education,* first issued in 1950. It at once became a standard text-book for students in training colleges. During the next few years it was reprinted eight times, and a fresh edition, thoroughly revised, was issued in 1960. His last book, *The Experimental Psychology of Beauty* (1962), had the same title as his first, but was extended to nearly 450 pages.

Among professional educationists and psychologists Prof. Valentine's name will always be associated with the *British Journal of Educational Psychology,* which he founded in 1931 and edited so efficiently for 25 years—a remarkable testimony to the old-fashioned policy of sticking to a good editor once you have found him and he has started to learn his complicated job. The new journal replaced the *Forum of Education,* which he had also founded, and had edited for nearly ten years. Like all pioneer journals, it at first encountered financial difficulties and various other frustrations. Nor would it have survived, had not Valentine himself acted as publisher, and managed its finances with little help except the gratuitous labour of his own family. When he resigned the editorship it was in a more flourishing condition than any of the psychological journals published in this country.

Throughout his academic life he was an active member of the Society and for many years one of the most valuable members of the Council. He was elected President for 1947–48. And here I should like to record a deep personal debt of my own—one among many. Valentine was one of the first to propose that the Society should establish a statistical journal, and to persuade Council to set aside a small sum to launch it. The journal itself was started while he was President; and in making the necessary editorial plans and arrangements I owed much to his untiring advice and support.

But most of all his name deserves to be remembered and honoured because, more than any other psychologist of his generation, he sought, deliberately and explicitly, to keep alive the tradition of British psychology, which, when he was a student, was 'more highly rated throughout the world than that of any other country'. He stood firm against the later fashion 'to look first across the Atlantic for every guiding idea'. As he observed in one of the last of his contributions at an Annual General Meeting, 'British psychologists of the present generation are re-learning from America views and conclusions which may seem new in that country, but which Stout, McDougall, and Myers were preaching half a century ago'; and, if one looks at the index of names or the list of references that concludes any one of his articles or books, one finds that the vast majority are English. Yet wide as has been the influence of his writings, it cannot possibly equal that which was exerted by his generous and vivid personality on all with whom he came into touch—students, colleagues, teachers, and educational officials throughout the country. There have been few in our lifetime to whom British psychology and British education owe so much.

C. BURT

Teachers sometimes regard writers on educational psychology as dreamers and armchair speculators, though many an overworked school teacher must have longed for the chance to drop occasionally into an armchair to ponder on the up-to-dateness, even the social significance, of subjects which the distant universities require him, sometimes against the promptings of his own judgement, to expound to helpless children. Prof. Valentine was no spinner of 'gray theory': in his youth he had been a schoolmaster, for most of his adult life he was an active university teacher, a seasoned and perceptive committee member, a keen researcher, with a fine grasp on reality when choosing themes to be investigated and techniques to be employed. Most of his researches were obviously useful, and for this fact he was not the man to offer justifications to austere-minded academics.. He was particularly interested in problems concerning children's mental development, studying his own family over a period of nearly twenty years.

To day it might not seem surprising if a psychologist were elected to a Chair of Education in Great Britain, but Valentine's appointments to Belfast and Birmingham were events to be recorded in history. Early in his career, he was interested, like many of his psychologist colleagues just then, in several questions, at that time defiantly begged, concerning the possibility of the 'transfer of training', or 'formal discipline'; a phrase which covered square-bashing, assiduous polishing of military equipment, to be hastily dulled on active service, and mechanical memorising of irregular verbs in a dead language. Valentine's experience as a teacher of Latin was valuable in writing his book *Latin, Its Place and Value in Education*. In it he examined aspects of this disputed subject; the efficacy of Latin as an instrument for intellectual training. At that time, many teachers, including some influential and articulate ones, believed that the mind consists of separate 'faculties'; that these, like muscles, can be automatically strengthened by formal training. Some teachers even conceived the school to be a mental gymnasium: an educationist commented that so too was Fagin's kitchen. The energetic reaction against the 'formal training' theory issuing in the researches of Thorndike in America and Sleight in England and their enthusiastic collaborators had been pushed too far, as Valentine and the *Gestalt* psychologists demonstrated. Transfer could take place, but in what conditions and with what vehicles? Scepticism was natural in any psychologist who at school had spent many hours of unenthusiastic work on Latin, very few on spoken French and none on German, spoken or written. Valentine inquired into the precise value of Latin to secondary school pupils who know from the start that they are to go no farther than a School Certificate standard.

His greatest contribution to psychology and education is probably the series of 25 volumes of the *British Journal of Educational Psychology*, which soon gained a world-wide reputation for original and trustworthy articles. Not only was he a very careful and a times a ruthless editor, but he 'went after' shy, reluctant or lazy writers who he knew had something valuable to say. In such choices he was not swayed by fashion, though schools of psychology were rapidly growing up, with rigid authoritarian discipline and ready to apply the classical defence of the hollow square. Nor did he disdain an occasional glance behind him, as when he commissioned by personal persuasion, a symposium 'Is the Doctrine of Instincts Dead?' This, far from being a nostalgic stirring of embers, reminded readers—and contributors—that though psychology had acquired a number of new terms for allegedly innate motives, it had not seen the last of instincts; an impression which the new ethologists soon deepened.

Though Valentine probably enjoyed an occasional fling into iconoclastic activity, he never ignored or condemned a belief because it was popular. He took seriously, and investigated carefully, the assumption that experienced teachers, especially women, could make valid intuitive judgements of a child after only a few minutes' oral examination or interview. The qualities to be assessed included not only general intelligence but also character traits such as temper, conscientiousness, obstinacy and the like.

Before we met, I had heard appreciations of Valentine at Würzburg. Külpe had recently been the host of three successful ambassadors for Britain. Two were from Oxford; in those days this might have seemed odd until one remembered that William McDougall was still there: the third, predictably, was from C. S. Myers at Cambridge. They were Cyril Burt, J. C. Flugel and C. W. Valentine. Külpe spoke warmly of them to me. I believe he regarded Valentine as typically English, carrying his learning with apparent lightness, and 'a great comic'. But Valentine could sound, if not look, serious enough if the occasion seemed to demand it. Soon after arrival he had learned, beside the necessary phrases for travellers, a few with special relevance. Asked why he had cut a certain lecture course, he would reply '*Nicht nützlich*!' Perhaps this too, properly appreciated was a comic statement. He loved reverberating German words, and was delighted to pronounce his address, thundering 'Randersackererstrasse!'

It was pleasant to work with him in examining, for his practice was as sound as his theory. Our tasks were never chores: for a lengthy consultation about borderline marks he characteristically chose a cool glade in Hyde Park, of course near the refreshment rooms. Whether examiners ought to write marginal comments on scripts is frequently discussed, but when he noticed an unconventional remark on psychologists by a bright or dim candidate, he would pen, in red ink, an unquantifiable remark, perhaps a quotation from a ribald poem, which had bubbled up from a nearby layer of his preconscious. I trust that these adornments lightened the tedium of the clerks who scrutinized the marked scripts, and that they were able to interpret some of the allusions.

I remember Valentine best from our meetings at conferences: the last in a happy six-some when the Valentines, the Wolters and the Pears found themselves living in the same hotel at Brighton. I enjoyed his quick appreciation of the comic and the ludicrous, which contributed valuably to his psychological commonsense. Has it not been claimed that one generation's commonsense is the 'science' of its predecessor?

It is a pleasure, in these days of kitchen-sink and ash-bin plays, of self-hating novelists and (many) self-ignoring psychologists, to write this tribute to an old friend who loved music, beauty and laughter. T. H. PEAR

Brit. J. Psychol. (1964), 55, 1, *pp.* 391–402
Printed in Great Britain

INTROJECTION: A FORM OF SOCIAL LEARNING*

By MICHAEL ARGYLE
Institute of Experimental Psychology, University of Oxford

Evidence from previous studies is interpreted in terms of introjection as the learning of self-reactions from the reaction of others. This is a form of learning which differs from others in that it is more permanent, generalizes to all situations, and exists in conflict with the rest of the personality.

A number of investigations are reported showing that introjection takes place for some children, but not others, particularly from the same-sexed parent, and when there is a warm and dependent relationship.

Several attempts to explain introjection in simpler terms are considered; it is concluded that the most plausible accounts are that parents persuade and teach children to react to their own behaviour in certain ways, and that two-person interaction may lead to fantasied anticipation of the other's reactions.

Introduction

The hypothesis which is to be presented is based on ideas first put forward by Freud, in his theory of the super-ego. According to this theory the super-ego is formed principally at the end of the Oedipal phase at about 4–5, when the child is forced to abandon the opposite-sex parent as a love-object. According to Freud's previously developed theory of mourning (1917), an abandoned object-choice is 'introjected'; that is to say, some parts of the object are ascribed to or copied by the self. It is deduced (1923) that the opposite-sex parent will now be introjected, and it is supposed that it is the attitudes of approval and disapproval for various acts which are introjected. Furthermore, this introjection does not fuse with the ego, merely producing a modification of behaviour: it forms a separate subsystem of the personality which reacts to the subject's behaviour in the same way as the parent did.

This theory contains a number of difficulties, of which the most important is its implication that the super-ego is based in some way on the opposite-sexed parent. Freud concluded, on the evidence of his own case-studies, that in fact the super-ego is based primarily on the demands made by the *same*-sexed parent, and this conclusion was confirmed by later studies (see below).

What can usefully be done is to restate the more plausible aspects of the theory. Let us postulate that when a person A is in a certain relation with a person in authority B, A will come to learn B's reactions to A's behaviour. Let us postulate that this learning forms a separate precipitate in A's personality—a set of anticipations of B's reactions. Guilt feelings will then be the introjections of B's withdrawal of love or approval; self-aggression will be the introjection of B's punishment. The internalized standards of the super-ego will be based on B's standards for A's behaviour.

The theory as developed so far does not state the conditions under which introjection occurs. It is shown elsewhere that some processes of social learning are not universally applicable, but take place only under certain conditions (Argyle & Delin, 1965). We shall review below previous investigations, and report some new ones,

---

* This research was supported by the Medical Research Council, which gave a grant for research assistance.

which establish between them the critical variables governing introjection. This will throw further light on the process itself.

PREVIOUS RESEARCH

Perhaps the most convincing demonstrations of the operation of introjection are the reported observations of individual young children who address themselves as their parents had done previously. Wälder (1937) reports on 2-year-old children who repeated parental demands in an imitation of the parental voice when faced with a temptation to disobey them. Sears, Maccoby & Levin (1957) report a similar case of a 17-month-old child.

Cooley's theory of the 'looking-glass self' (1902) can be regarded as a special case of introjection—where *A* comes to adopt *B*'s cognitions about himself. There is a certain amount of evidence from small-group studies that self-perceptions correlate with the perceptions of others (Zetterberg, 1957), but this could be due to perceptual agreement rather than to the process of introjection. Racial minority groups sometimes come to accept the evaluation of themselves by the dominant group, as in the 'self-hatred' of some American Jews, and the approval of light-coloured skins among negroes (Allport, 1954). An experiment by Staines (1958) showed that after 12 weeks of teaching in a way designed to improve the self-pictures of children, self-ratings were significantly more favourable. Further support comes from studies of professional training, where it is found for example that medical students who are treated as doctors by families for which they are responsible come to see themselves as doctors (Merton, Reader & Kendall, 1957). Similarly, a case is reported in which a group of male students decided to treat a rather dull and unattractive girl as if she was tremendously attractive and popular. 'Before the year was out she had developed an easy manner and a confident assumption that she was popular' (Guthrie, 1938). Chinese thought reform (a 'brain-washing' technique making use of group pressures) is successful in getting people to accept the group's evaluation of them as guilty criminals (Lifton, 1957).

Further evidence comes from studies of the socialization origins of guilt and self-aggression. Two predictions can be made: that withdrawal of love techniques will result in guilt feelings, and that physical punishment will result in self-aggression. The first hypothesis follows from the introjection hypothesis if it is granted that guilt consists of withdrawing love or approval from the self. A large number of studies have found a small positive correlation between withdrawal of love and guilt (Hoffman, 1963). Self-aggression is partly due to the punishment and subsequent displacement of outward aggression, as is shown by the correlation found between it and punishment for aggression, in several studies (e.g. Sears, 1961). Is it also due to introjection of physical punishment? Sears (1961) found a correlation of 0·22 for boys ($P < 0{\cdot}06$), 0·10 for girls (N.S.) between total punishment at age 5, and a questionnaire measure of self-aggression at 12. Funkenstein King & Drolette (1957) found that subjects who directed aggression inwards perceived their fathers as the main source of authority in the family, and unlike the subjects who directed aggression outwards they had a close and affectionate relationship with them. It is likely that such authoritative fathers would have used physical punishment. This is not to overlook the fact that the usual result of physical punishment is outward aggression on the part of the child; what the above evidence suggests is that under certain conditions there may be self-aggression as well as, or instead of, the outward aggression.

Several studies suggest that introjection is a form of learning that has certain special features. (1) Introjection results in the acquisition of response tendencies which are to some extent in conflict with other parts of the personality. This is most clearly seen in the case of the conscience, which is commonly experienced as coming partly from 'outside the self'. Clinical studies of authoritarians show that their consciences are particularly externalized in this way (Adorno *et al.* 1950). People feel that they 'ought' to act in a certain way, not that they 'want' to. In the case of the changes of self-perception discussed earlier there is some degree of acceptance of the cognitions of others. There is evidence at least in the early stages that the new self-perceptions are only partially accepted, and co-exist with other self-perceptions. Kelman (1961) has argued that these changes in the case of thought reform are brought about by the immediate relationship with the influencing agents and fade away at the end of this relationship.

(2) When introjection takes place there is an internalization of the responses involved, so that they occur not only in the presence of the person who was the original source of influence, but generalize to all situations. Research into the origins of the conscience distinguishes a stage of public conformity at about 3–4 when children obey parental demands in the presence of the parents; at 5–6 many children internalize these demands and obey them in the absence of the parents (Leedham, 1956). Not all people do form a conscience in this way—those who are 'other-directed' rather than 'inner-directed' (Riesman, 1950), or who are controlled by 'shame' rather than guilt (Piers & Singer, 1953) clearly do not.

(3) The effects of introjection seem to be permanent and do not decay with time. The evidence for this is that the introjections of parental demands made at the age of 5 are detectable in adults, and are indeed found by psychoanalysts particularly difficult to modify. This may be due to the isolation of introjected systems in the personality, which insulates them from the impact of later modifying experiences. It is this permanence of early introjections which accounts for the fact that people feel most guilty about the trivial misdemeanours of early childhood, rather than about motoring and financial offences.

Finally, several studies throw some light on the conditions under which introjection does and does not occur. Sears *et al.* (1957) found a correlation of only 0·09 between withdrawal of love discipline and conscience-formation in 379 5-year-olds. However, for the warm mothers 42 % of the love-withdrawal children had strong consciences as compared with 24 % of those who had experienced little or no love-withdrawal. For the cold mothers there was no such difference. McCord, McCord & Zola (1959) found a considerably lower rate of delinquency for boys whose fathers used psychological discipline, but again only when there was a warm relationship. Miller & Swanson (1960) found that psychological discipline was related to indirectness of expression of aggression for middle class children only.

For a strong conscience to develop there must be a warm relationship, and strict discipline. It has been found that mental patients have been reared with greater use of psychological discipline, and less use of physical punishment, than have controls

(Greenfield, 1959). It is often found that neurotically disturbed children, particularly those who are anxious and guilty, come from homes where the discipline is very strict (Hewitt & Jenkins, 1946). Anxious and depressive states can be regarded as in part due to an over-severe super-ego which differs from that of the authoritarian in that it *is* internalized. The background for this appears to be an intense relation with the parents, together with strictness and psychological discipline.

Most delinquents, on the other hand, seem to have weak super-egos in that they not only break the rules, but are deficient in guilt feelings. The family background of many delinquents is one of rejection or neglect by parents, and discipline which is lax or over-severe. Clearly none of the conditions for introjection occur, so that neither guilt, nor self-aggression, nor any other form of restraint is acquired.

EMPIRICAL STUDIES OF INTROJECTION

(1). *Introjection and achievement motivation*

The investigation to be described was reported previously (Argyle & Robinson, 1962) so it will be summarized briefly. Projective and questionnaire measures of need for achievement were given to several samples of grammar-school children. The children described the behaviour of their parents, using the semantic differential. The finding relevant to our present hypothesis is that need for achievement ('n.Ach') correlated significantly with reported demands for achievement on the part of parents. This was predicted from the introjection hypothesis—children should introject the parental demands and thus have a high n.Ach. This interpretation is confirmed by the fact that estimates of super-ego achievement tendencies, 'how hard I ought to work' correlated with the projective measure of n.Ach, and with reported parental demands. It was found that the correlation between n.Ach and parental demands was stronger for those children who had a warm relation with their parents, or who identified with them.

A similar result has been obtained by Rosen & d'Andrade (1959), who found that boys with high n.Ach had parents who set high standards for them.

(2) *Individual studies using Q-technique*

Q-technique has been used before for obtaining statistical tests of hypotheses for individuals. For example, Welford (1947) asked subjects to rank-order six situations for the extent to which these would be experienced as frustrating, and the likelihood of the subject praying in them. The average rank-order correlation was 0·59, but this was much higher for some subjects than others, while for some there was very little relation. In the present investigation rankings were made both by children and by their parents to obtain evidence about the effects of socialization techniques on individual children.

Hypothesis. If parents react to a child's behaviour in a certain way, the child will come to react in the same way to himself. In this study four types of parental response to misbehaviour were considered, and it was predicted that they would produce four kinds of self-reaction. The four sequences are as follows:

(1) Parent laughs at child—child feels silly.
(2) Parent indicates child has failed—child feels sense of failure.

(3) Parent is shocked, withdraws love—child feels guilty.

(4) Parent uses physical punishment—child directs aggression to itself.

Method. The subjects were twenty 9- to 10-year-old children, ten boys and ten girls, and their parents. The children were selected so that all were normal and had good relations with their parents.

The Q-sort consisted of sixteen cards on each of which was typed some way of being naughty—'not washing', 'being greedy', 'breaking things in the house', and so on. The cards were lettered from A to P.

Table 1. *Individual correlations between self-reactions and parental discipline*

	Mother–daughter	Father–son
All subjects: average of twenty individual correlations ($N = 16$ for each)		
Failure	0·22**	0·17**
Guilt	0·19**	0·22**
Self-aggression	0·06	0·21**
Feels silly	0·20**	0·05
Introjectors: average of six individual correlations ($N = 16$ for each)		
Failure	0·51†	0·54†
Guilt	0·49†	0·27†
Self-aggression	0·09	0·23
Feels silly	0·15	0·11

** $P < 0·01$.

† For an average of six correlations, each with $N = 16$, $P < 0·01$ when $r = 0·26$, but since these subjects were specially selected this level of probability cannot be used.

The child was asked to pick out and put in rank order: (1) 'the things that would make you feel silly if you did them'. When he had completed this, the child read out the letters on the cards, and the order was noted by E. The child then sorted the cards three more times for 'the things that make would you feel (2) guilty, or bad inside, (3) that you have failed in some way, or ought to have done better, (4) so cross with yourself that you could kick yourself'.

The parents were tested individually and asked to put the cards in order to show the extent to which they had reacted by: (1) 'laughing at the child', (2) 'showing that you are hurt or shocked', (3) 'indicating that the child has failed or fallen short of a standard', (4) 'spanking the child'.

Results. The four sortings by the child and those by the two parents were converted to rank orders and all the correlations computed for each child separately.

The first hypothesis was that self-reactions would be related to discipline experienced. For twenty subjects, in four areas, the average correlation for the same-sexed parent was 0·17: an average of 80 correlations with $n = 16$ in each case ($P < 0·001$). The average correlation between other combinations of conscience and discipline, an average of 240 correlations, was not significantly different from zero.

As is shown in Table 1 the hypothesis was confirmed most strongly for failure and guilt, and for six children out of the twenty (three of each sex) in particular, who showed the effect in three or four areas. It was stronger for the same- than for the opposite-sex parent, and only the former will be reported here.

Discussion. This study shows that introjection works on average for these twenty subjects over four areas and both parents, but that it works in particular for six of those subjects, for the same-sexed parent, and for two of those areas (failure and guilt).

(3) *Correlates of introjection*

The next step with the Q-correlations is to find out the socialization variables with which they in turn are correlated. This has not yet been done, but some other correlates have been found. It was expected that the more parental demands had been introjected, the stronger the general level of conscience would be. A questionnaire was constructed in which subjects were asked to rank ten personal attributes for THE KIND OF PERSON I OUGHT TO BE and the KIND OF PERSON MY FATHER (MOTHER, FRIENDS) THINK I OUGHT TO BE. Rank-order correlations were worked out as before.

For the eighty-four students to whom this was administered, guilt was not related to introjection of parent's demands, but it was correlated with introjection of friends' demands. The correlation was higher for male students ($r = 0{\cdot}32$, $P < 0{\cdot}05$), and the average for the whole sample was 0·23 ($P < 0{\cdot}05$).

For other samples the semantic differential was used to obtain the distance between the two sets of perceived demands (see below). This was an unsatisfactory measure, since subjects used the extreme points of the scales, so that for example a large difference score mainly reflected the perception of parental demands as feeble. However, for thirty-four theological students it was found that introjection of parental demands correlated with guilt at $r = 0{\cdot}33$ ($P < 0{\cdot}05$). For the other students guilt was not correlated with introjection of parents' demands but was correlated with introjection of friends' demands.

(4) *Studies with the semantic differential*

It follows from the introjection hypothesis that children should feel they ought to do or be whatever their parents want them to do or be. Thus the detailed contents of the super-ego should be closely modelled on parental demands.

The contents of the super-ego were explored by means of the Semantic Differential (Osgood, Suci & Tannenbaum 1957). Subjects filled in a number of 7-point rating scales to indicate THE SORT OF PERSON I OUGHT TO BE. Then they completed a similar set of scales to indicate THE SORT OF PERSON MY FATHER THINKS I OUGHT TO BE, etc. The similarity between two such profiles can be computed by one of several alternative 'distance' scores. We found that the various scores correlated highly together and we used the simplest to compute—the arithmetical mean of differences of scale positions for all the scales. The subjects used were eighty students and 500 grammar-school children.

For 160 boys and male students the super-ego was more similar to father's demands than to mother's demands ($P < 0{\cdot}0015$), and to father's demands than to father's reported behaviour ($P < 0{\cdot}0003$). For the girls, there was a non-significant tendency for the super-ego to be nearer mother's demands than father's. For the school children the super-ego was nearer to parental demands than to friends' demands.

The semantic differential scales were grouped into four factors, and scores were worked out for these. It was found that the super-ego factor scores correlated with corresponding parental demand factor scores in the range 0·35–0·70. Other groups of subjects were given *Q*-sorts to describe super-ego and parental demands; the average rank-order correlation was again about 0·50. These results provide further evidence for the introjection of specific parental exhortations.

It may be objected that these associations between super-ego and reported parental demands could be due to similar response sets operating, since the word 'ought' appears in both sets of instructions. This could not of course account for the differential effects found for the two parents, or for the comparison between parental demands and friends' demands.

(5) *Studies of degrees of externalization*

Our hypothesis was that Catholics should have a more externalized super-ego than Protestants—in view of the greater authoritarianism of Catholics (Adorno *et al.* 1950), and the less close parent–child relationships which would be expected to accompany the greater use of physical punishment (Argyle, 1958). Externalization of the super-ego was measured by reported difference in descriptions of MYSELF AS I ACTUALLY AM and THE KIND OF PERSON I OUGHT TO BE. When *Q*-sorts are used, a lower rank order between the two sortings is taken to indicate conflict and externalization.

Table 2. *Self/super-ego conflict and religious denomination*

The figures represent average correlations between self and super-ego rankings. A correlation was worked out for each subject; these are the averages for the number of subjects given. The difference between the two marked '*a*' is significant at $P < 0.02$.

	R.C.	*N*	C. of E.	*N*	Non-Conf.	*N*	No Church	*N*
Males	0·17	37	0·24	85	0·22	24	0·22	85
Females	0·10*a*	36	0·23*a*	97	0·23	25	0·19	41

It can be seen from Table 2 that the hypothesis is strongly confirmed, especially for females. Other analyses of the data show that there is more externalization of introjections when there are many siblings ($r = 0.25$, $P < 0.01$), and low identification with parents (males only, $r = 0.41$, $P < 0.001$).

SUMMARY OF THE CRITICAL VARIABLES GOVERNING INTROJECTION

The studies reviewed earlier, and those reported above, show in a number of different ways that introjection in the sense defined can occur under certain conditions. As the individual studies using *Q*-technique show, it is far from being a universal process, but one that occurs with a minority of children. We shall now review briefly the evidence on what the special conditions are under which it does work.

Warmth of relationship between parent and child was shown to be essential in studies by Sears *et al.* (1957) and McCord *et al.* (1959) and was shown to be important by Funkenstein *et al.* (1957) and Argyle & Robinson (1962).

Sex of parent. The super-ego is influenced most by the same-sexed parent. Blum (1949) gave his Blacky test to students; in one card subjects were asked to identify the sex of a big dog threatening a toy dog; subjects generally supposed the big dog had the same sex as themselves. Andry (1960) found that rejection by the father was more closely linked with delinquency in boys than was rejection by the mother. Sears *et al.* (1957) found that rejection by father weakened conscience formation in boys but not in girls. Sears (1961) obtained some evidence that self-aggression in boys was particularly connected with the relationship with father. Funkenstein *et al.* (1957) found that self-aggression occurred in male students when there was a close relation with a powerful father. The importance of the same-sex parent was confirmed in the Q studies reported above (2) and in the semantic differential studies (4). It is likely, however, that where the reactions of the opposite sex are important, as in the incident reported by Guthrie, opposite-sex reactions can have greater effect.

Dependence was found to be a factor in conscience-formation by Sears *et al.* (1957), and may be a necessary condition for introjection. The very fact that introjection has been found mainly in the parent–child relationship suggests that this is so. Where members of groups introject group reactions, they are usually subordinate low-status members, and introjection probably only occurs here if there is an intense and dependent relation, as in Chinese thought reform.

Age. As reported above, children as young as 18 or 24 months can be observed speaking to themselves in the parental voice. However, permanent introjections are not formed until 5–6, and there is probably some upper limit after which parental exhortations cease to be introjected. Group reactions can be introjected in the case of young adults, under the conditions mentioned above.

ATTEMPTS TO EXPLAIN INTROJECTION IN TERMS OF SIMPLER PROCESSES

(1) *Avoidance conditioning*

Various theories have been put forward by Eysenck (1957) and Mowrer (1950, 1960) to account for the super-ego in terms of avoidance conditioning. The basic idea is that experience of punishment for carrying out a certain act leads to anxiety or guilt about this, together with the acquisition of behavioural restraints. The parental discipline is conceptualized in terms of punishment or the withholding of reward, and the super-ego as anxiety plus avoidance learning.

Several research findings are directly relevant to this theory. First, it follows that the more children are punished the stronger should be the super-ego. As far as physical punishment is concerned, the opposite is found (Sears *et al.* 1957). The role of physical punishment seems to be restricted to controlling behaviour in the presence of the punitive agent, but has little carry-over to other situations. Boys who are beaten for being aggressive at home are less aggressive at home but more aggressive at school (Sears *et al.* 1953). Numerous studies show that delinquents have received *more* physical punishment than others. It is withdrawal of love which seems to produce true super-ego formation: if this can be interpreted as a more severe kind of punishment than physical punishment the theory could stand, but there is nothing within the theory to suggest why this should be so. Secondly, if we grant that introverts have a stronger super-ego than extroverts, the finding that introverts condition better gives some support to the present theory (Franks, 1961). It is suggested by Eysenck (1957) that introverts are innately predisposed to condition better, and hence acquire stronger consciences in childhood.

However, the theory is quite unable to explain the sequence of events which we are calling introjection, and the same is true of all attempts to explain super-ego formation in terms of schedules of punishment. Furthermore, this theory is unable to explain special features of this kind of learning—its generality and its separateness.

(2) *Identification*

Sears *et al.* (1957) and Whiting & Child (1953) have put forward an explanation in terms of identification. It is postulated that the young child identifies with his parents, and imitates their behaviour towards him. This is reinforced both by their approval for such imitation, and by the fact that the behaviour reproduced includes approving and nurturant actions, which the child now fills in for himself in the parents' absence. If this is to be explained in terms of instrumental learning by reward, it is hard to see why the characteristically self-critical self-punishing aspects of the super-ego should be learnt—it would rather be expected that these responses would be extinguished, while the self-rewarding ones should be retained.

This theory does, however, offer an explanation of the importance of love-oriented techniques of discipline. If identification is reinforced by parental love, it follows that techniques will be most effective which 'are likely to have the dual effect of keeping the child oriented toward the goal of parental affection and at the same time arousing uncertainty about the attainment of this goal' (Whiting & Child, 1953, p. 242).

Secondly, a warm relationship with the parents should be connected with super-ego development, on the grounds that this indicates reward for identification. The deduction seems dubious, but the prediction is confirmed. Psychological methods of discipline are found to be only effective when there is warmth—suggesting that the true role of warmth is in making psychological discipline effective.

Thirdly, a dependent relationship with the parents should be an antecedent of conscience, on the grounds that this indicates a strong drive for parental affection. Sears *et al.* confirmed this prediction, though the effect was greater for boys than girls; Whiting & Child failed to confirm it.

The fact that girls have the stronger super-ego is explained by reference to the shift in identification from mother to father which boys make, and some evidence is provided for such a shift (Sears *et al.* 1957, p. 384). It could have been added that since the mother provided early nurturance there would be more love to be lost for girls. The importance of the same-sexed parent is explained, given that identification is usually with the same-sexed parent. Perhaps the most distinctive prediction from this theory is that children with strong consciences should identify with their parents. No evidence has been reported to this effect, and a series of studies by the author has failed to find any such correlation. The theory does account for the withdrawal of love–guilt sequence, but not for other types of introjection.

(3) *Identification with the aggressor*

According to this theory the child engages in *defensive* identification with a harsh and punitive parent. Sanford (1955) stresses that this is essentially a reaction to a crisis and that it produces a rigid and unrealistic caricature of the parental behaviour; he suggests that the parent is not fully introjected into the ego, and gives the externalized super-ego of the authoritarian personality as a typical instance. There should then be identification, of a special sort, with parents who are very punitive. Anna Freud (1946) reported individual cases of such children, and Bettelheim (1943) observed that some concentration camp prisoners identified with guards, slavishly copying their ill-treatment of other prisoners. However, socialization research shows that children identify more with parents who are warm and nurturant. Mowrer (1950) suggested that there could be two kinds of identification—developmental and defensive, based on nurturant and punitive relationships respectively. Bronson (1959) studied 42 boys aged 9–13, half of whom had a stressful relationship with their fathers; while some of the latter rejected their fathers as models, others displayed an exaggerated masculinity and caricatured their fathers. This suggests that defensive identification with the aggressor can occur with *some* people under the conditions prescribed. The theory could account for the introjection of physical punishment. Again the theory requires that introjection should correlate with identification, which is not generally true. It could not explain the origins of guilt.

(4) *Introjection as the learning of self-reactions*

The self-reactions of which introjection consists could be learnt by normal learning processes. Perhaps parents can teach children to make certain self-reactions, as is suggested by the use of such expressions as 'you ought to be ashamed of yourself' and 'can't you see how silly it is to do that'. Parents can exercise sheer persuasion and then give reinforcing approval if the child reacts in the right way. A child may also discover that making the appropriate self-reactions will reduce parental punishment. Thus learning may occur by cognitive restructuring, reward or avoidance. Since parents want children to react in the same way as they do, introjection should take place. It is a little difficult to see how physical punishment could be introjected, unless a parent adopts the attitude 'don't you think that a person who does that deserves to be spanked' and succeeds in persuading the child to agree. Aronfreed, Cutick & Fagen (1963) suggest that self-reactions can be learned as avoidance reactions, as they terminate the anxiety due to anticipated punishment.

From this formulation we would expect introjection to occur, not necessarily when the child identifies with the parent, but when it is open to influence, as would be the case with a warm and dependent relationship. It does not explain the greater influence of the same-sexed parent. It does account for the separateness of the super-ego, as a bundle of self-reactions somewhat in conflict with other tendencies.

Two investigations provide more direct evidence about this theory. Aronfreed (1963) found that children who had been experimentally trained to deprive themselves of material rewards for deviant responses were more likely to make reparations after an experimentally encouraged transgression. Bandura & Walters (1963) found that inhibited, guilty boys, compared with aggressive, non-guilty boys, reacted to situations of deviance by describing the character in a story as exhibiting guilt and trying to make restitution. 'The occurrence of a deviation was a cue for self-criticism and for attempts to reinstate positive reinforcers through the making of amends.' In addition the evidence on modification of self-perceptions, in thought reform and elsewhere, shows how self-cognitions may be acquired through such social learning.

(5) *Role introjection*

This theory is based on an analysis of the social interactions between the introjector (A) and the influencing agent (B). It is assumed that A will modify his behaviour in accordance with B's reactions, especially if B is powerful and if A is dependent on B. A comes to learn how B will react to whatever A does; he learns to anticipate B's responses. It is now postulated that A may enact B's responses in fantasy in the absence of B. Boisen (1936) described the process of conscience-formation in these terms: 'In language and imagination we rehearse the responses of others just as we dramatically enact other consequences. We foreknow how others will act, and the foreknowledge is the beginning of judgement passed on action.'

One prediction is that when A is good at role-taking he should introject more readily. There is some confirmation for this from the finding that delinquents (who have weak consciences) are poor at role-taking, i.e. at enacting or predicting the reactions of other people (Sarbin, 1954). A second prediction is that introjection should occur when A is dependent on B, which was confirmed by Sears *et al.* (1957). Thirdly, the theory predicts that introjection will occur when the social relationship has been of a certain degree of intensity—as it is with parents, and in the case of thought reform.

This theory looks hopeful but it depends on further speculations about two-person interaction, and in the absence of more knowledge about it does not take us very far forward.

I am grateful to Peter Delin and the late Mrs Adrienne Dunn, who worked as research assistants on this project.

REFERENCES

Adorno, T. W. *et al.* (1950). *The Authoritarian Personality.* New York: Harper.

Allport, G. W. (1954). *The Nature of Prejudice.* Cambridge, Mass.: Addison Wesley.

Andry, R. G. (1960). *Delinquency and Parental Pathology.* London: Cape.

Argyle, M. (1958). *Religious Behaviour.* London: Routledge and Kegan Paul.

Argyle, M. & Delin, P. (1965). Non-universal laws of socialization. *Hum. Relat.* (in press).

Argyle, M. & Robinson, P. (1962). Two origins of achievement motivation. *Brit. J. Soc. Clin. Psychol.* **1**, 107–20.

Aronfreed, J. (1963). The effects of experimental socialization paradigms upon two moral responses to transgression. *J. Abnorm. Soc. Psychol.* **66**, 437–48.

Aronfreed, J., Cutick, R. A. & Fagen, S. A. (1963). Cognitive structure, punishment, and nurturance in the experimental induction of self-criticism. *Child Develpm.* **34**, 281–94.

Bandura, A. & Walters, R. H. (1963). *The Social Learning of Deviant Behavior.* New York: Holt, Rinehart and Winston.

Bettelheim, B. (1943). Individual and mass behavior in extreme situations. *J. Abnorm. Soc. Psychol.* **36**, 417–52.

Blum, G. S. (1949). A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.* **39**, 3–99.

Boisen, A. T. (1936). *The Exploration of the Inner World.* New York: Harper.

Bronson, W. C. (1959). Dimensions of ego and infantile identification. *J. Personality,* **29**, 532–45.

Cooley, C. H. (1902). *Human Nature and the Social Order.* New York: Scribner.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

FRANKS, C. M. (1961). Conditioning and abnormal behaviour. In Eysenck, H. J. (ed.). *Handbook of Abnormal Psychology*. London: Pitman.

FREUD, A. (1946). *The Ego and the Mechanism of Defense*. London: Hogarth Press.

FREUD, S. (1917). Mourning and melancholia. *Collected Papers*, IV, 152–70.

FREUD, S. (1923). *The Ego and the Id*. London: Hogarth Press. (Standard edition, vol. 19.)

FUNKENSTEIN, D. H., KING, S. H. & DROLETTE, M. (1957). *Mastery of Stress*. Harvard University Press.

GREENFIELD, N. S. (1959). The relationship between recalled forms of childhood discipline and psychopathology. *J. Consult. Psychol.* **23**, 139–42.

GUTHRIE, E. R. (1938). *The Psychology of Human Conflict*. New York: Harper.

HEWITT, L. F. & JENKINS, R. L. (1946). *Fundamental Patterns of Maladjustment, the Dynamics of their Origin*. Illinois: D. H. Green.

HOFFMAN, M. L. (1963). Childrearing practices and moral development: generalization from empirical research. *Child Developm.* **34**, 295–318.

KELMAN, H. C. (1961). Processes of opinion change. *Pub. Op. Quart.* **25**, 57–78.

LEEDHAM, L. R. (1956). An empirical investigation into areas of moral awareness. M.A. thesis, University of British Columbia.

LIFTON, R. G. (1957). Methods of forceful indoctrination. In Stein, M. R. *et al.* (eds.), *Identity and Anxiety*. Glencoe, Illinois: Free Press.

McCORD, W., McCORD, J. & ZOLA, I. (1959). *The Origins of Crime*. New York: Columbia University Press.

MERTON, R. K., READER, G. & KENDALL, P. L. (1957). *The Student-Physician*. Harvard University Press.

MILLER, D. R. & SWANSON, G. E. (1960). *Inner Conflict and Defense*. New York: Holt.

MOWRER, O. H. (1950). *Learning Theory and Personality Dynamics*. New York: Ronald.

MOWRER, O. H. (1960). *Learning Theory and the Symbolic Processes*. New York: Wiley.

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press.

PIERS, G. & SINGER, M. B. (1953). *Shame and Guilt*. Springfield, Illinois: Thomas.

RIESMAN, D. (1950). *The Lonely Crowd*. Yale University Press.

ROSEN, B. C. & D'ANDRADE, R. (1959). The psychosocial origins of achievement motivation. *Sociometry*, **22**, 185–218.

SANFORD, N. (1955). The dynamics of identification. *Psychol. Rev.* **72**, 106–18.

SARBIN, T. R. (1954). Role theory. Chap. 6, in Lindzey, G. (ed.), *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley.

SEARS, R. R. (1961). Relation of early socialization experiences to aggression in middle childhood. *J. Abnorm. Soc. Psychol.* **73**, 466–92.

SEARS, R. R. *et al.* (1953). Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.* **49**, 135–234.

SEARS, R. R., MACCOBY, E. E. & LEVIN, H. (1957). *Patterns of Child Rearing*. Evanston, Illinois: Row, Peterson.

STAINES, J. W. (1958). The development of children's values. III. The self-picture as a factor in the classroom. *Brit. J. Educ. Psychol.* **28**, 97–111.

WÄLDER, R. (1937). The problem of the genesis of psychical conflict in earliest infancy. *Int. J. Psycho-Anal.* **18**, 406–73.

WELFORD, A. T. (1947). Is religious behavior dependent upon affect or frustration? *J. Abnorm. Soc. Psychol.* **42**, 310–19.

WHITING, J. W. M. & CHILD, I. L. (1953). *Child Training and Personality*. Yale University Press.

ZETTERBERG, H. L. (1957). Complaint actions. *Acta sociologica*, **2**, no. 4.

(*Manuscript received* 15 *November* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 403–414
Printed in Great Britain

IGNORING IRRELEVANT INFORMATION

By P. M. A. RABBITT
Medical Research Council Applied Psychology Research Unit, Cambridge

The task of scanning a visual display to identify certain symbols may involve ignoring the presence of others, and so may be compared with Donders's 1862 study of the '*c*-reaction', in which a response was required to only one of a set of five possible phonemes presented to a subject in random succession. For visual displays of letters of the alphabet the time required to ignore a symbol as irrelevant is shown to depend upon the vocabulary of items for which search is conducted. Incidental learning of irrelevant symbols is shown to be related to the same variable. It is shown that performance in situations of this type is analogous to that in which subjects respond to groups of stimuli rather than to particular items, and stimulus and response entropy are independently varied.

Introduction

One hundred years ago Donders (1862) undertook a line of investigation which has since been neglected, to the cost of our understanding of perceptual processes. Donders presented members of a set of five phonemes to his subjects on successive trials, requiring them to respond to one, but not to the others ('*c*-reaction'). Mean response times to the critical phoneme were found to be 46 msec. longer than in a situation in which it was presented as the only stimulus in a simple-reaction task ('*a*-reaction'). It is thus clear that subjects require a measurable time to ignore any member of a class of similar stimuli. The effects of variations in the numbers of relevant and irrelevant stimuli used in this situation have never been examined, perhaps because attention has been concentrated upon three other approaches to the problem of discrimination between relevant and irrelevant material.

(1) *Concept formation studies*

The identification of stimuli as members of categories such as 'relevant' and 'irrelevant' has been a topic for learning theory and 'concept-formation' studies. In such studies the typical assumption has been that in everyday life class responses such as 'cow' or 'brick' are made by selecting some cues as critical and ignoring others as irrelevant or fortuitous variation. The problem of ignoring irrelevancies has therefore been treated as that of discrimination between critical and redundant aspects of complex displays. In this tradition Archer (1954), Gregg (1954) and Hodge (1959) have investigated the time taken to respond to some aspects of complex displays, such as shape, when other aspects or dimensions of difference such as colour, position and size must be ignored. The same line of thinking has been followed by Baker (1961) when assessing the effects of visual 'noise' on discrimination of signals on radar displays. Donders's situation represents a special case of this argument in which the relevant and irrelevant stimuli vary along the same continua.

(2) *Selection between sense-organs*

Cherry (1954), Moray (1959) and Treisman (1960) have examined the effects of competing auditory stimulation upon the perception of speech. Subjects were required to select between competing monologs, repeating aloud whichever was designated by the experimenter. Such studies have greatly improved our understanding of selection processes, and are in many ways logically parallel to Donders's study. However, although the techniques used by these authors are very elegant, they impose two serious limitations on discussion of their data:

(i) The number of different relevant and irrelevant items to be distinguished has never been an independent variable. The methods of quantification which can be applied to these results are therefore quite limited.

(ii) It is not easily possible to study the time characteristics of the selection process by such means; consequently we do not know how long it takes to ignore an irrelevant item.

(3) *Data from eye-movement recording*

Technical advances in eye-movement recording have enabled Mackworth & Mackworth (1958) to describe the temporal sequence in which complex visual displays are scanned. Estimates were thus obtained for the time required to identify items under a wide variety of conditions of stimulus complexity. Grace, Warhurst & Murdock (1962), using simpler techniques, have also investigated the time taken to locate groups of letters and numbers of varying complexity in displays containing similar material. In the wide variety of experiments of which these cases are representative, subjects have always been required to search for single items specified by the experimenter in advance. These studies therefore represent exact visual analogues of Donders's situation, with the drawback that the conditions of the experiments made it impossible to estimate the time required to discriminate between relevant and irrelevant stimuli. The further extension of Donders's experiment in which the number of relevant and irrelevant items are independently varied has not been undertaken in these or other terms. Experiment I (which was carried out in 1959) undertakes this for a situation in which subjects search for complex stimuli (letters of the alphabet) on displays in which other similar stimuli must be ignored.

Experiment I

Effects of increases in vocabulary of relevant symbols upon the time taken to ignore irrelevant symbols

Material

Eight packs of forty-eight visiting card blanks measuring 2¼ in. × 3½ in. were made up to vary the number of irrelevant symbols present in a display under two conditions of response entropy:

Condition I. The first four packs offered a 2-choice task, half the cards in each pack being stencilled with the letter A and half with the letter B. Letters were block capitals 5 mm. high, stencilled in indian ink. These letters appeared equally often at each of nine possible locations on a card as shown in Fig. 1. Subjects sorted packs so as to separate cards marked with A from cards marked with B.

The cards gave four conditions of irrelevancy:

Pack 1. No other letter appeared on the cards.

Pack 2. One other letter of the alphabet appeared on each card. These letters were of the same size and type-face as the relevant letters, all letters of the alphabet from C to Z being used equally often for this purpose. Positions of irrelevant letters were randomized among the eight unfilled locations.

Pack 3. Four irrelevant letters from the same vocabulary were stencilled on each card in the same way.

Pack 4. Eight irrelevant letters from the same vocabulary were stencilled on each card so that all nine positions were occupied and the display resembled that in Fig. 1.

Condition II. Packs 5–8 offered an 8-choice task, each card being stencilled with one of the letters A, B, C, D, E, F, G, or H. The four conditions of irrelevancy used in condition I were duplicated using the letters I–Z.

It will be noted that the range of relevant + irrelevant letters was 26 in both conditions of this experiment.

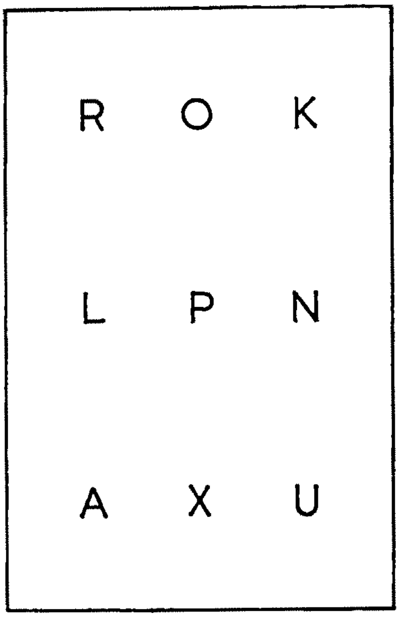


Fig. 1. Card with nine locations filled.

Procedure

Subjects sorted packs of cards, holding them face-downward in the non-dominant hand and turning them for inspection before placing them in the appropriate compartment of a sorting-frame. The sequence of stimuli in each pack was randomized by shuffling before it was sorted. Times were taken with a stopwatch.

Nine subjects began with Condition I, sorting each of the packs 1–4 ten times in random order. They then attempted Condition II, receiving the same amount of practice. The remaining nine subjects reversed the order in which they experienced the conditions.

The results presented below are based on the means for the last two trials on each pack.

Subjects

The eighteen subjects used were four female shop assistants aged 17–31 (mean age 22·4), four male and four female undergraduates aged 18–24 (mean age 20·3) and six Royal Navy ratings aged 19–30 (mean age 25·2).

Results

Results are given in Table 1. Analysis of variance was carried out on the data with the conclusions given below:

Condition I. Differences in sorting-time between Pack 1 (0 irrelevant letters) and Pack 2 (1 irrelevant letter) are significant, as are those between Pack 2 (1 irrelevant letter) and Pack 3 (4 irrelevant letters) ($P < 0{\cdot}01$). The difference between Pack 3 and Pack 4 (8 irrelevant) is not significant ($P > 0{\cdot}1$).

Table 1. *Mean sorting time/card in seconds as a function of number of irrelevant symbols on display*

Number of irrelevant symbols in display	0	1	4	8
Condition I (2-choice)	0·36	0·52	0·60	0·62
Condition II (8-choice)	0·65	0·81	0·90	1·11

Condition II. All differences between packs are significant ($P < 0{\cdot}01$).

The overall difference between Condition I and Condition II was significant ($P < 0{\cdot}001$). The interaction term for sorting time × conditions × number of irrelevant letters on the display was significant ($P < 0{\cdot}05$). Because the slope for the increase in response time with number of irrelevant letters is almost identical for the first three packs in each condition, the significant interaction term must be attributed to the plateau occurring at the transition between Pack 3 and Pack 4 in Condition I.

Consideration of results of Experiment I

In Condition I response-time rises as the number of irrelevant symbols on a display is increased from 0 to 4 (Packs 1–3) but not when it is further increased from 4 to 8 (Packs 3 and 4). Some feature of the display must therefore counteract the effects of an increase in the number of irrelevant symbols between these limits. Data from eye-movement studies (Mackworth & Mackworth, 1958) suggest that subjects do not fixate words or symbols upon a display one at a time in succession, but rather that they group items perceptually, processing several with a single fixation. It follows that such grouping becomes easier as the density of items on a display increases, and the mean distance between them is accordingly reduced. In the present experiment irrelevant and relevant symbols were widely scattered over the displays in Packs 1–3. It is therefore probable that subjects had to make as many independent fixations to scan displays in Pack 3 (4 irrelevant) as in Pack 4 (8 irrelevant). The benefits of greater symbol density in Pack 4 and the effects of an increase in the number of symbols upon the display would therefore tend to cancel each other out.

Condition II shows no such discontinuity in the effects of increases in the number of items in a display upon response-time. This suggests that, in this case, subjects sample symbols as single items, or at all events in smaller groups than in Condition I. The most important inference to be drawn from Expt. I is therefore that the number of symbols which can be simultaneously processed as a group is not constant, and is inversely related to the number of items for which search is conducted.

The question as to whether subjects take longer to ignore an irrelevant symbol in

Condition II (8 relevant) than in Condition I (2 relevant) is oversimple, since this is seen to depend upon the time required to process a 'group' of symbols, and more particularly upon the limiting number for such a group in each condition. The mean distance between symbols was not held constant in the displays used in Expt. I. The fact that increases in response-time with transitions from 0 to 1 and from 1 to 4 irrelevant symbols are almost identical in Conditions I and II may simply indicate that in the first three packs of each condition grouping is limited by the mean distance between symbols on a display rather than by the number of items for which subjects are set to search. It was therefore necessary to repeat the experiment with displays in which the mean distance between symbols was held constant. This was done in Expt. II below.

EXPERIMENT II

Repeat of Experiment I with displays on which the mean distance between symbols remains constant

Material

Sixteen packs of forty-eight cards each were made up, eight being used for each of two conditions of the experiment:

Condition I. The eight packs of this condition offered a 2-choice task, cards marked with the letter B as described in Expt. I. In Pack 1 of this condition the relevant letter was stencilled in the centre of each card, and no other letters were present. In Pack 2 the relevant letter was accompanied by one other letter of the alphabet, drawn from the vocabulary of letters C–Z. Irrelevant letters were stencilled in the same size, ink and face as the relevant letters. The irrelevant letter appeared equally often before and after the relevant letter, so as to form, with it, a nonsense digram: e.g. AK, OB, etc. In Packs 3–8 the number of irrelevant letters was progressively increased from two (Pack 3) to seven (Pack 8). The position of the relevant letter in the nonsense words so formed was systematically varied from card to card within each pack so that, as far as the arithmetic of the situation allowed, the relevant letter appeared equally often in each possible position. The spacing between letters was kept constant in all packs.

Condition II. The eight packs of this condition offered an 8-choice sorting task, five of the forty-eight cards in each pack bearing one of the relevant letters A, B, C, D, E, F, G, or H. The conditions of irrelevancy, and controls for embedding relevant among irrelevant letters precisely duplicated those described for Condition I.

It will be noted that in both conditions of this experiment, as in both conditions of Expt. I, the vocabulary of relevant + irrelevant letters was kept constant at 26.

Procedure

Subjects sorted these packs into the appropriate number of compartments of a sorting-frame. The procedure for administering and timing was precisely as described for Expt. I. Half the subjects experienced the conditions in one order and half in the other. Subjects sorted each pack in each condition 10 times. Packs within each condition were presented in random order. The results given below are based on the means for the last two trials given each subject on each pack.

Subjects

Twenty-four Royal Navy ratings aged between 17·8 and 31·0 years (mean age 22·4) served as subjects.

Results

Results are given in Table 2.

The outcome of analysis of variance on the data is described for each condition separately, and then for interactions between conditions.

Condition I. The increase in sorting time with the number of irrelevant letters present upon a display is significant over all 8 packs in this condition ($P < 0{\cdot}001$). This increase appears to be stepwise rather than continuous, since examination of differences between the means for individual packs by t-tests shows that the sorting times for Pack 2 (1 irrelevant), Pack 3 (2 irrelevant) and Pack 4 (3 irrelevant) are not significantly different. Sorting times for Packs 5 (4 irrelevant) and Pack 6 (5 irrelevant) are also not significantly different. All other differences between packs are significant ($P < 0{\cdot}01$).

Table 2. *Mean sorting-time/card as a function of the number of irrelevant symbols on the display*

Number of irrelevant symbols on the display	0	1	2	3	4	5	6	7
Condition I (2-choice)	0·54	0·65	0·66	0·64	0·68	0·68	0·69	0·70
Condition II (8-choice)	0·86	0·93	0·97	1·01	1·05	1·10	1·13	1·18

Condition II. The overall increase in sorting time with the number of irrelevant symbols upon a display is significant ($P < 0{\cdot}001$). All differences between individual packs are significant by t-test ($P < 0{\cdot}01$).

Interaction between conditions. The interaction term for Conditions × Sorting time × Number of irrelevant symbols is significant ($P < 0{\cdot}01$). It must be concluded that the effect of increases in the number of irrelevant letters in a display upon sorting time is greater in an 8-choice than in a 2-choice task.

Consideration of results of Experiment II

Condition I again shows discontinuities in the effects of the number of irrelevant symbols in a display upon sorting time. This confirms the inference drawn from Expt. I that, in certain cases, the effects of increases in the number of irrelevant symbols are masked because symbols may be processed in groups of more than one at a time. In this case it might be expected that the effects of increases in the number of irrelevant symbols would only be noticed when the number of relevant + irrelevant symbols exceeds a maximum 'group size', or when symbols are spatially so far apart that they must be scanned one at a time by separate fixations. The present data do not, unfortunately, allow us to estimate the maximum group size in either condition, and it can only be said that the group-size is smaller in Condition II than in Condition I. The time required to locate a symbol buried in a nonsense-word of irrelevant symbols may be expected to vary with its position in that word (i.e. first, second, ..., *n*th ... last letter). In Expt. II the positions of relevant letters in the nonsense words made up of relevant + irrelevant letters were systematically varied within each pack. As a result, the sorting time for each pack represents the summed response-times to as many different display conditions as there are relevant + irrelevant letters on the cards. The precise variations of 'group size' with the vocabulary of relevant symbols employed, and the time required to process such groups can only be determined by obtaining individual response times to a range of displays of these types.

Since it clearly takes longer to decide that a symbol is irrelevant when searching for as many as eight possible items in a display than when searching for either of two, it is a reasonable inference that the perceptual analysis of all letters in the display, relevant and irrelevant, is more complicated in the 8-relevant case. Fitts *et al.* (1956) have shown that when learning to discriminate between complex stimuli, such as the patterns they employed, subjects gradually discovered which cues, or points of difference between patterns were critical, and which were misleading or redundant and therefore must be ignored. The number of points of difference which must be learnt in order to discriminate between patterns will increase both with the complexity of the patterns and with their number.

Letters of the alphabet are also complex stimuli. The efficiency with which they are discriminated from each other will therefore also depend upon the degree to which subjects are able to select among critical cues. Such selectivity will also be limited by two independent factors:

(i) By the number of different items to be discriminated from each other.

(ii) By the physical characteristics of the particular subset of symbols within which these discriminations must be made. Cues which are optimal when discriminating between one subset of symbols may be misleading or redundant when discriminating between members of another subset.

The degree to which there is transfer from practice on discriminations between members of one subset to discriminations within a different subset will thus depend on the degree of identity of the cues sampled in the two cases. The more cues are involved, the more radical will be the readjustments necessary, and consequently also the more re-learning the subject will have to do. Subjects may be expected to require fewer cues in order to discriminate two relevant symbols from each other, and from a group of irrelevant symbols, than to discriminate eight relevant symbols from each other, and from a group of irrelevant symbols. It should then follow that fewer readjustments to the cue system should be necessary on transfer to a new group of irrelevant symbols in a 2-choice than in an 8-choice situation. Expt. III tests this hypothesis.

EXPERIMENT III

Comparison of cue-sampling in 2-choice and 8-choice tasks

Material

Four packs of forty-eight cards were made up, two for each condition of this experiment:

Condition I, 2-choice task. Cards were prepared as for Pack 4, Condition I, Expt. I (cf. Fig. 1), each card being stencilled with either the letter A or the letter Z in one of nine possible locations. The remaining eight locations on each card were occupied by irrelevant letters.

In Pack 1 the irrelevant letters were J, Y, Q, B, H, P, C and L.

In Pack 2 the irrelevant letters were T, K, F, N, R, S, D and U.

Condition II, 8-choice task. Cards were prepared as for Pack 8, Condition II, Expt. I, each card being stencilled in one of nine locations with one of the eight letters A–H. The remaining locations on each card were occupied by irrelevant letters. In Pack 1 the irrelevant letters were I, J, K, L, M, N, O and P. In Pack 2 the irrelevant letters were Q, R, S, T, U, V, W and X.

In all these packs the positions of relevant and irrelevant letters were randomized between the nine possible positions on each card as described for Pack 4, Expt. I.

Procedure

Subjects were given one of the four packs, told which were the relevant letters, and required to sort the cards as rapidly as possible into the appropriate number of piles (2 or 8). Procedure for sorting and timing were as described for Expts. I and II. Subjects were given five successive runs with the pack with which they had begun, and were given knowledge of their sorting times on each occasion. On the sixth trial subjects were required to sort the second pack of the same condition, the substitution being made without comment. The sorting time obtained on this trial was used as a criterion of the extent of transfer to a situation in which a different set of irrelevant symbols had to be ignored.

On completing one condition of the experiment subjects were given a brief rest, and were then given the remaining condition in the same way. Half the subjects began with Condition I and half with Condition II. Each of the sixteen subjects employed sorted the packs in a different one of the twenty-four orders possible under these conditions.

Subjects

Twelve male undergraduates aged 20–22 years (mean age 21·6) and four Royal Navy ratings aged 17–26 (mean age 23·8) were used as subjects.

Results

The mean sorting times per card for the last trial on the practice pack and the first trial on the transfer pack in each condition are given in Table 3.

Transfer to a pack with a new set of irrelevant symbols significantly increases response time in both conditions ($P < 0{\cdot}01$ on analysis of variance).

Table 3. *Mean sorting-time/card for practice and transfer packs, and practice:transfer ratios for 2- and 8-choice conditions*

	Condition I, 2-choice task (sec.)	Condition II, 8-choice task (sec.)
Practice pack	0·71	1·33
Transfer pack	0·82	1·69
Ratio, practice:transfer	0·87	0·79

In order to test the relative effects of transfer in a 2-choice and in an 8-choice situation, the ratios of practice to transfer were obtained for each subject in both conditions and compared by a Wilcoxon matched pairs signed ranks test. The ratios obtained for the 8-choice condition were significantly smaller than those obtained for the 2-choice condition ($P < 0{\cdot}01$, two-tailed equivalent) showing that it is more difficult to adapt to a change in the vocabulary of irrelevant symbols in the 8-choice than in the 2-choice task.

DISCUSSION AND CONCLUSIONS

It follows from Expts. I and II that the speed with which relevant and irrelevant symbols are discriminated from each other depends upon the number of symbols for which the subject has to search. These data do not fit any simple information-theory model, since although the *number* of irrelevant symbols appearing upon the cards was increased from pack to pack within each condition, the total *vocabulary* of symbols, relevant and irrelevant, remained constant at twenty-six for all packs in both conditions. The information value of any given letter on a display was thus identical in 2-choice and 8-choice conditions.

It cannot therefore be the case that subjects identify every letter on a display in turn, making the appropriate decision as to where to place a card when they arrive at a relevant letter. If this were so it should take no longer to identify irrelevant letters in an 8-choice situation than in a 2-choice situation, since the vocabulary of irrelevant letters is actually smaller in the 8-choice case (18) than in the 2-choice case (24).

We must conclude that the irrelevant letters are not discriminated from each other. Relevant letters must, of course, be discriminated both from each other and from irrelevant letters. In this context the task of ignoring irrelevant letters may be compared to that of making the same response to all members of an arbitrary group of stimuli. A 2-choice task would therefore be regarded as a case in which subjects differentiate between three classes of stimuli: the class of the letter A, the class of the letter B and the class of all irrelevant letters C–Z. An 8-choice task can be regarded as a case in which subjects respond to nine classes of stimuli, the classes of A, B, C, D, E, F, G, H and the class of irrelevant letters I–Z.

Rabbitt (1959) and Pollack (1963) have shown, in different categorization tasks, that the stimulus vocabulary and the number of classes into which these stimuli are categorized by subjects' responses interact in their effect on response time. That is to say, increases in the total vocabulary of stimuli which do not affect response time in a 2-choice task will result in sharp increases in response time in an 8-choice task. The finding that subjects take longer to categorize symbols as irrelevant in a 9-class categorization task (8 relevant + irrelevant category) than in a 3-class categorization task (2 relevant + irrelevant category) is thus consonant with earlier work on categorization situations, of which the task of ignoring irrelevant stimuli may be regarded as a special case.

Expt. III further suggests that improvement in the ability to discriminate between symbols may be related to the extent to which subjects learn specific sets of cues which are optimal for discriminations within the set with which they are practised. When transfer takes place such cues may no longer be valid for the new discriminations required. Subjects must then re-learn new cues, and performance is impaired until this is done. Subjects were found to be relatively more slowed by transfer to a new set of irrelevant symbols in an 8-choice task than in a 2-choice task. Since the relevant symbols were the same in both the practice and transfer conditions, we must assume that the cues to be re-learnt are necessary to discriminate relevant from irrelevant symbols. It follows that subjects must master a wider range of cues in the 8-choice than in the 2-choice task. To put this somewhat differently, we may say that subjects learn more about the irrelevant symbols when searching for any of eight relevant symbols than when searching for either of two.

It is important to distinguish the implications of these experiments from those of three other approaches to problems of discrimination:

(i) *Information theory*

(*a*) It has been emphasized that matching processes must concern physical points of difference between stimuli. The number of cues necessary to distinguish between stimuli will tend to increase with the number of stimuli to be discriminated. However, it is theoretically possible to construct equal numbers of patterns in such a way that in order to discriminate between members of one set twice as many cues must be considered as would be necessary to distinguish members of the other. Similarly, it might be expected that discrimination of the letter O would be less affected by the addition of angular letters such as A, H, M or X to a display than by the addition of rounded letters such as C or Q. The complexity of standards required in order to make discriminations of this type may thus be independent of the information load of the task as calculated from the number of items to be discriminated. A way must be found to apply the information metric to the vocabulary of *cues* rather than to the vocabulary of *stimuli* when stimuli are other than points along a single continuum of operation of a sense-organ.

(*b*) The grouping of stimuli into 'relevant' and 'irrelevant' categories appears to control the efficiency with which subjects can distinguish between items in a display. The limitations of the system involved are clearly not a function of information load in any simple sense. Variations in the vocabulary of *irrelevant* symbols (from twenty-four in the 2-relevant situations to eighteen in the 8-relevant situations) are clearly much less critical than variations in the number of relevant stimuli in determining the time taken for discriminations between relevant and irrelevant categories. The results obtained by Pollack (1963) and Rabbitt (1959) may best be summarized by the statement that the time taken to choose between categories of stimuli varies as a multiplicative function of the stimulus population and the number of categories into which this population is divided by the subject's responses. Both Pollack and Rabbitt investigated only situations in which all classes or categories always contained the same number of stimuli. The present experiments may best be treated as situations in which subjects respond to classes of stimuli of unequal sizes: i.e. two or eight classes of 1-relevant stimulus per class, and a class of eighteen or twenty-four items. The nature of the functions obtained under these conditions, and their relationship to the expressions derived from information theory which have so far been applied to choice response situations, require further investigation.

(ii) *Weber's Law*

It must be emphasized that the term 'discriminability' is not used as in cases to which the Weber–Fechner function and quantification in j.n.d.'s are applicable. We are concerned here with the *range* of cues necessary to discriminate between complex stimuli, and the selectivity which can be exercised in using some and ignoring others. The *obviousness* of such cues may be estimated in j.n.d.'s or some equivalent system of quantification, and is an important consideration not covered by these data. The implications of the present experiments are thus different from those which follow from the work of Crossman (1955) and Rabbitt (1963), namely that there is an interaction between the discriminability of stimuli, as measured in j.n.d.'s along some continuum of difference, and their information load.

It is very probable that variations in the discriminability of stimuli, in the sense brilliantly discussed by Crossman, will affect the efficiency with which an optimal set of cues may be selected. Such effects will represent the operation of yet another variable in this complex situation. In such terms as have been employed above, this distinction may be paraphrased as that between the *range* of cues which must be incorporated in standards used for recognition and the *tolerance* with which such cues may be matched.

(iii) *Span of absolute judgement*

When a subject is required to identify a stimulus presented to him in the absence of other members of a set to which the stimulus belongs, he is said to make an absolute judgement. Experiments concerned with discriminations of this type have typically been conducted with stimuli representing points along a single continuum of operation of a sense-organ. That is to say, subjects have been required to discriminate between tones differing in loudness but not in frequency, or to make identifying responses to a series of positions of a point along a single spatial continuum. Miller (1956), reviewing the literature in this area, argues that human beings do not appear to be able to learn to discriminate between more than a few points along any continuum of operation of a sense-organ. Miller remarks that the number of such absolutely discriminable points (7 ± 2) is curiously similar for all sensory continua for which such data are available. Pollack (1953) has further shown that such limitations are not necessarily related to the size of intervals along the sensory continuum between these points—in other words to their mutual separation in j.n.d.'s. Such a distinction precisely parallels that made in (i) above between the discriminability of stimuli and their information load.

Complex stimuli, with which the present discussion is concerned, differ from each other in terms of more than one sensory continuum. As Miller (1956) points out, little is known about the span of absolute judgement for multi-dimensional stimuli. Experiments by Klemmer & Frick (1953) on the span of apprehension for positions of a dot on one and two spatial dimensions, and by Pollack & Ficks (1954) on absolute judgement for tones differing in both frequency and intensity, indicate that the span of absolute judgement increases with the number of dimensions of difference between stimuli, but is not a simple multiplicative function of the number of dimensions of difference and the number of steps discriminable along each dimension in isolation. However, caution is necessary before these findings are generalized, since although there are about seven discriminable points along a single spatial continuum and about thirty-six discriminable points within a square (Klemmer & Frick, 1953), it does not follow that there would be a similar reduction of the number of discriminable steps along each continuum of difference if subjects were required to identify points by their *colour* and *position*. This is a further difficulty in the application of the information theory metric to these situations, since it is probable that there are interactions between the numbers of discriminable steps along some pairs of continua and not others. The further investigation of this problem appears to offer a first approach to the understanding of perceptual coding.

If matching of perceptual input to some internal standard is held to be the basis for recognition processes it must also follow that there is a limit to the number of distinct states of input that can be modelled by the brain as standards of comparison. This question is basic to the understanding of 'set' and becomes particularly interesting when we consider how it may be that a 'category' of many distinct stimuli can be modelled or matched in this way.

While the present data do not allow us to do more than raise these problems, we may note in conclusion that more detailed description of the mechanism by which we identify classes of items will provide quantitative predictions on the operation of perceptual set, or as Külpe (1904) remarked, 'knowing what to look for'. It is a source of pleasure to the present author that this will only be made possible by a return to the realism of this great investigator; that is, to a refusal to consider discriminations between complex stimuli without giving thought to the ways in which their physical characteristics are selected and modelled by the brain.

REFERENCES

ARCHER, R. J. (1954). Identification of visual patterns as a function of information load. *J. Exp. Psychol.* **48**, 313–17.

BAKER, C. H. (1961). 'Man and Radar Displays'. Prepared for 'The Advisory group for Aeronautical Research and Development', N.A.T.O. at Defence Research Board, Department of National Defence, Canada.

CHERRY, E. C. & TAYLOR, W. K. (1954). Some further experiments upon the recognition of speech with one and with two ears. *J. Acoust. Soc. Amer.* **26**, no. 4, 554–9.

CROSSMAN, E. R. F. W. (1955). The measurement of discriminability. *Quart. J. Exp. Psychol.* **7**, 176–95.

DONDERS, F. C. (1862). Die Schnelligkeit psychischer Processe. *Arch. Anat. Physiolog.* **1862**, pp. 657–82.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, N. & LEONARD, J. A. (1956). Stimulus correlates of visual pattern recognition. *J. Exp. Psychol.* **51**, 1–11.

GRACE, M. R., WARHURST, F. & MURDOCK, B. B. (1962). Search time for alpha-numeric stimuli. Research memorandum prepared under contract 3219 (02) (x) at the Department of Psychology, University of Vermont, Burlington, Vermont, U.S.A.

GREGG, L. W. (1954). The effect of stimulus complexity on discriminative responses. *J. Exp. Psychol.* **48**, 289–95.

HODGE, M. H. (1959). The influence of irrelevant information upon complex visual discrimination. *J. Exp. Psychol.* **57**, 1–5.

KLEMMER, E. T. & FRICK, F. C. (1953). Assimilation of information from dot and matrix patterns. *J. Exp. Psychol.* **45**, 15–19.

KÜLPE, O. (1904). Versuche über Abstraktionen. *K. Exp. Psychol., Berl.*, **1**, 56–68.

MACKWORTH, N. H. & MACKWORTH, J. F. (1958). Eye fixations recorded on changing visual scenes by the television eye-marker. *J. Opt. Soc. Amer.* **48**, 7, 439–45.

MILLER, G. (1956). The Magical Number 7 ± 2. *Psychol. Rev.* **63**, 81–97.

MORAY, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quart. J. Exp. Psychol.* **11**, 56–60.

POLLACK, I. (1953). The information of elementary acoustic displays. II. *J. Acoust. Soc. Amer.* **25**, 765–9.

POLLACK, I. (1963). Speed of classification of words into superordinate categories. *J. Verb. Learning Verb. Behav.* **2**, 159–65.

RABBITT, P. M. A. (1959). Effects of independent variations in stimulus and response probability. *Nature, Lond.*, **183**, 1212.

RABBITT, P. M. A. (1962). Perceptual discrimination and the choice of responses. Unpubl. Ph.D. thesis, University of Cambridge.

RABBITT, P. M. A. (1963). Stimulus information load and discriminability. *Nature, Lond.*, **197**, 726.

TREISMAN, A. M. (1960). Contextual cues in selective listening. *Quart. J. Exp. Psychol.* **12**, 242–8.

(*Manuscript received* 18 *January* 1963; *Revised* 28 *March* 1964)

Brit. J. Psychol. (1964), 55, 4, *pp.* 415–419
Printed in Great Britain

SEQUENTIAL ERRORS IN A TIME-SHARING TASK

By G. S. TUNE
Harvard University

An experiment is described in which subjects were required to maintain a set pattern of responses while simultaneously engaged in a monitoring task. The results showed that selective deterioration in performance took place which was demonstrable only by means of a sequential analysis of responses and not by a simple count of the errors made. The results are discussed in relation to a concept of performance decrement which proposes that 'regression' through the various levels at which a task is organized is a function of the capacity of short-term memory.

I. Introduction

Welford (1953) has argued that the hierarchical organization of behaviour is inevitably dependent on the capacity of short-term memory. If such a capacity is reduced then behaviour will become disorganized, because, to use Welford's terminology, the control of elements 'higher' in the scale over those lower down will be loosened. A simple example of this is the following. If an operator's task consists of predicting which of two possible events (1 and 0) will next occur, and they always occur in the invariant order 11001100..., the operator can invariably predict a future event with complete accuracy if he is able to remember the two previous events. Thus if he knows that 01 has just occurred, he may predict the next event with complete certainty. On the other hand, if the operator can only remember *one* previous event, his prediction of what the next event will be cannot be made with such certainty. In this situation, when the necessary capacity of short-term memory is lacking, operators are to be expected to exhibit disorganized behaviour, where possibly the correct responses are given but in the wrong order. In other words, the sequential characteristics of the responses represented by the transitional probabilities ($p(11)$, $p(10)$, etc.) would not reflect the characteristics of the events, but the simple probabilities of occurrences of the responses ($p(1)$ and $p(0)$) might. In the double alternation situation (110011001100...) if operators were predicting all events with complete accuracy not only would they have 100% correct responses, but the responses would share the same probability characteristics as the events, namely, $p(11) = p(10) = p(01) = p(00) = 0{\cdot}25$; $p(1) = p(0) = 0{\cdot}5$. Under conditions of reduced short-term memory capacity one of two situations might occur:

(1) The responses might all be wrong, or inappropriate, but still maintain the correct probability characteristics, such as when a subject gave responses which were consistently out of step, e.g.

```
     EVENTS 1 1 0 0 1 1 0 0 1 1 0 0
            * * * * * * * * * * * * <- incorrect responses
PREDICTIONS 0 0 1 1 0 0 1 1 0 0 1 1
```

This would give zero correct responses, because all the predictions were inappropriate, but a correct set of probability characteristics.

---

[1] Now at the Medical Research Council Unit, University of Liverpool.

(2) The predictions might be partly inappropriate and also have incorrect (or partly incorrect) probability characteristics, e.g.

EVENTS	1	1	0	0	1	1	0	0	1	1	0	0	
		*	*			*	*			*	*		← incorrect responses
PREDICTIONS	1	0	1	0	1	0	1	0	1	0	1	0	

This would give a correct response score above zero and below 100 % but only partly correct probability characteristics, i.e.

$$p(11) = 0, \quad p(10) = 0{\cdot}5, \quad p(01) = 0{\cdot}5, \quad p(00) = 0: \quad p(1) = p(0) = 0{\cdot}5.$$

Obviously there are many sets of transitional probabilities which will yield the desired probabilities of occurrence, e.g.

$$\left.\begin{array}{l} p(11) = 0{\cdot}1 \\ p(10) = 0{\cdot}4 \end{array}\right\} p(1) = 0{\cdot}5 \qquad \left.\begin{array}{l} p(11) = 0{\cdot}3 \\ p(10) = 0{\cdot}2 \end{array}\right\} p(1) = 0{\cdot}5$$

$$\left.\begin{array}{l} p(10) = 0{\cdot}4 \\ p(00) = 0{\cdot}1 \end{array}\right\} p(0) = 0{\cdot}5 \qquad \left.\begin{array}{l} p(01) = 0{\cdot}2 \\ p(00) = 0{\cdot}3 \end{array}\right\} p(0) = 0{\cdot}5$$

This formulation shows that the effect of an agent which reduces the capacity of short term-memory may be evaluated by (*a*) a simple count of the inappropriate responses that operators make, and (*b*) an examination of the sequential characteristics of the responses, preferably both. Because of a lack of research which reports both types of analysis an experiment was designed to examine whether an operator's responses to a learned and invariant visual task would be disordered in the manner outlined above as a result of time-sharing with an auditory task.

II. EXPERIMENT

Two visual stimulus event sequences were constructed; each was eight events long and each event was of the type R or G.

SEQUENCE 1. G R R R G G R R / G R R R G G R R...
SEQUENCE 2. R R R R R G R G / R R R R R G R G...

Apparatus and procedure

An event was the illumination of either a red (R) or a green (G) lamp; the lamps were placed on the subject's response panel immediately above two response keys. Pressing and releasing a response key brought on the next event and recorded the class of response made. The events were scheduled to appear in the order of either sequence 1 or sequence 2 by means of a system of uniselectors and wander plugs wired to a programming board. Subjects were used one at a time and sat before the response panel where they were instructed to guess or predict which of the two lamps would next come on. It was explained that predictions were to be registered by pressing and releasing the appropriate key and that knowledge of results was available by observation of the events themselves. No indication was given that the events would be presented in a recurrent order. Subjects continued predicting until they had learned the sequence to which they had been assigned and were able to repeat it three times without error. At this stage they were given earphones and told to continue to predict the occurrences of the lights *without error* and simultaneously listen to the auditory signals. The auditory stimuli were a set of randomized digits from 0 to 9 inclusive, recorded on tape at a speed of one digit per sec. in a monotonous voice. Subjects were asked to report hearing the digits 1 and 2 either singly or successively by pressing a third response key placed midway between the other two keys on the response panel. When it was ascertained that the subject had understood the instructions, the experimenter left the room and the experiment began with the first spoken digit. Each subject continued with his tasks until he had made 200 predictions.

Subjects

Twelve subjects were used; all were male University undergraduates between the ages of 18 and 24 and had no idea of the purpose of the experiment. Each subject was randomly allocated to one of two groups of six; each group experienced only one of the two visual sequences.

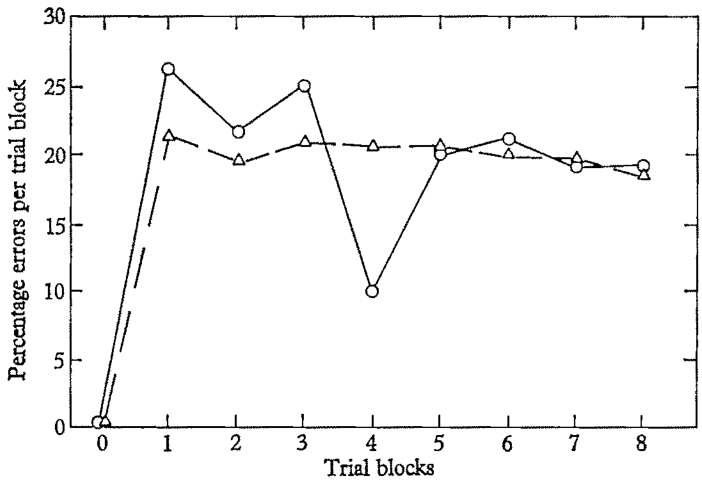


Fig. 1. The percentage of predictive errors is shown plotted as a function of trial blocks of twenty-four predictions. Each point is the mean error score of six subjects.

O—O, Sequence no. 1; △--△, sequence no. 2.

III. RESULTS

A block of twenty-four predictions was termed a trial (i.e. three repetitions of a sequence). The mean number of incorrect predictions made per trial was computed for both sequences and the results are plotted in Fig. 1 where each point is the mean of six subjects. It appeared that the number of errors in predicting the occurrences of the visual stimuli after the introduction of the auditory task was in the region of 20–30 %, with a slight improvement as the experiment proceded. Errors made in detecting the auditory signals were low (about 9 % over the whole experiment for both sequences). This analysis showed that in order to deal with both tasks errors were made. The second analysis was designed to examine the pattern taken by the errors in the visual task.

The time-ordered sequences of the subjects' responses to the visual task were listed and tabulated in overlapping blocks of three (thus the sequence R R R G G R would be tabulated as R R R, R R G, R G G, etc.). This gave a count of the frequencies with which subjects used each of the eight possible trigram patterns, from which the digram and engram frequencies were calculated. These data are shown in Table 1 along with a similar treatment of the event sequences, so that a comparison might be made between what was presented and what subjects thought was to be presented. For example, in sequence 1, 12·5 % of the trigrams presented were of the type R R R, but 15·4 % of the trigrams which subjects used were of this pattern; similarly 37·5 % of the digrams in sequence 1 were of the type R R, but 38·4 % of the digrams used by the subjects were of this type. This analysis gave an accurate estimate of the type of sequential errors made and allowed more information to be gathered about the subjects' performance than a simple count of the number of incorrect predictions made. The mean discrepancy between presented and reported trigrams in sequence 1 was 1·32 % and in sequence 2, 3·03 %. A Mann-Whitney U-test showed that the discrepancies in sequence 2 were significantly greater with $P < 0{\cdot}052$. A similar examination of the digram errors gave sequence 2 (mean error 5·43 %) a significantly greater error than sequence 1 (mean error 0·83 %) with $P < 0{\cdot}014$. The mean engram errors were 0 for sequence 1 and 4·5 % for sequence 2.

Table 1. *Results of the sequential analysis of sequences* 1 *and* 2 *showing the percentages with which the trigrams, digrams and engrams were presented and reported*

Trigram			Digram			Engram		
Pattern	Presented frequency (%)	Reported frequency (%)	Pattern	Presented frequency (%)	Reported frequency (%)	Pattern	Presented frequency (%)	Reported frequency (%)
Sequence no. 1								
R R R	12·5	15·4	R R	37·5	38·4	R	62·5	62·5
R R G	25·0	23·0						
R G R	12·5	11·8	R G	25·0	24·1			
R G G	12·5	12·3						
G R R	25·0	23·0	G R	25·0	24·2	G	37·5	37·5
G R G	0	1·2						
G G R	12·5	12·1	G G	12·5	13·3			
G G G	0	1·2						
Sequence no. 2								
R R R	37·5	47·0	R R	50·0	59·9	R	75·0	79·5
R R G	12·5	12·9						
R G R	25·0	18·7	R G	25·0	19·6			
R G G	0	0·9						
G R R	12·5	12·8	G R	25·0	19·5	G	25·0	20·5
G R G	12·5	6·7						
G G R	0	0·9	G G	0	1·0			
G G G	0	0·1						

It was anticipated that the distracting effects of the auditory task would reduce the subjects' capacity to predict accurately the occurrences of the visual events. The precise manner in which this could take place is described in the Introduction (p. 415) and the results of the sequential analysis mirrored this fairly well. In sequence 1, subjects generated exactly the correct proportions of 1's and 0's but they were predicted in the wrong orders, as shown by the discrepancies between the presented and reported digram and trigram frequencies in Table 1. In sequence 2, however, both the proportions of the two types of event predicted and the orders in which they were predicted were largely incorrect. Along with the results of the Mann-Whitney tests, this may be taken as an indication that the disorganization of sequence 2 was greater than that of sequence 1, a finding which was not apparent from a simple count of the number of incorrect predictions made.

IV. DISCUSSION AND CONCLUDING REMARKS

A task which requires the use of four previous events in the prediction of a fifth will not be performed adequately if the subject can remember only three of them; performance will be even more inadequate if only two or one of the necessary events can be used (it is not intended to imply by this that the process is conscious). In the present experiment the stimulus events were learned and occurred in an invariant order; it was possible, therefore, for the subjects to have predicted every event with complete accuracy. The results, however, showed that the introduction of the secondary, auditory monitoring, task brought about a selective breakdown in the performance of at least one of the sequences used. This can be interpreted in terms of Welford's thesis that a reduction in the span of short-term memory reduces the control of 'higher' hierarchical elements over the lower. Both stimulus sequences were complex in that the prediction of any event with complete accuracy required the knowledge of several previous events. The secondary task divided subjects' attention between the demands of the visual and auditory displays sufficiently to reduce the number of previous events which could be used as predictors and thus performance deteriorated both sequentially and in terms of the number of inappropriate predictions made.

There are three concluding comments, leading from this brief discussion, which may be made. First, the importance of events which serve as predictors is evident, as is the importance of their unavailability. In many tasks more mundane than those used in the laboratory, however, the exact predictors which operators use are unknown or hard to verbalize. The identification of these cues for action and the arrangement of them in such a way that the operator can take full advantage of them is a problem area which might be fruitful, especially for Ergonomics. Secondly, Schouten, Kalsbeek & Leopold (1962) have emphasized that the study of decaying performance can give insights into the formal organization of behaviour, which, once known, might provide valuable diagnostic tools for the clinical sciences. Finally, it is many years since Miller & Frick (1949) first drew attention to the fact that the sequential characteristics of behaviour were of interest and importance. Surprisingly little advantage has been taken of this potentially rich area of research and it is to be hoped, perhaps not too piously, that some of the problems which have been mentioned here will be further explored.

The work reported in this paper was carried out at the University of Bristol while the author held a D.S.I.R. award. The author is grateful to Stephen Griew and Peter Powesland for their encouragement.

REFERENCES

MILLER, G. A. & FRICK, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychol. Bull.* 57, 311–23.

SCHOUTEN, J. F., KALSBEEK, J. W. H. & LEOPOLD, F. F. (1962). On the evaluation of perceptual and mental load. *Ergonomics*, 5, 251–60.

WELFORD, A. T. (1953). The psychologist's problem in measuring fatigue. In Floyd, W. F. & Welford, A. T. (eds.). *A Symposium on Fatigue*, pp. 91–7. London: Lewis.

(*Manuscript received* 23 *July* 1963; *revised* 13 *January* 1964)

Brit. J. Psychol. (1964), 55, 4, *pp.* 421–427
Printed in Great Britain

RELATION BETWEEN PERCEPTUAL AND CONCEPTUAL OPERATIONS: NUMERICAL DISTANCE AND VISUAL EXTENT*

By L. TEFT, S. WAPNER, H. WERNER and J. H. McFARLAND
Clark University, Worcester, Mass., U.S.A.

This study deals with the problem of interaction of conceptual (number) and perceptual (extent) processes. The general hypothesis was tested that if an extent is demarcated by two numbers (from 1 to 9), the perception of extent varies depending on the numerical interval used. It was found that a visual extent demarcated by a smaller numerical interval appeared equal to a visual extent demarcated by a greater numerical interval when the former extent was physically smaller relative to the latter.

INTRODUCTION

This study deals with a problem area which has come more and more into the centre of perceptual theory, i.e. the relation between perceptual (sensorial) and conceptual operations. A crucial notion pertinent here has been that of 'schema'. This term was introduced by Kant to formulate the relation between concept and percept. In Kantian terms: on the one hand, the sensorial material is formed and organized into perceptual objects through the mediation of spatial and temporal schemata; on the other hand, spatial and temporal schemata are utilized in the concretization of the abstract concepts. Thus, schemata, for Kant, are mediating operations which internally link sensorial-perceptual and conceptual processes.

In recent times, the concept of the schema has been utilized as a principle for perceptual organization (cf. Head, 1926; Werner, 1940; Bartlett, 1950; Piaget, 1952; Vernon, 1955). Developmentally considered, objects are presumed to be formed primordially through sensory-motor schemata and later through internalized conceptual schemata. A conceptual schema pertinent to our work here is the number schema, by which number becomes defined in terms of a spatial order. Because spatialization plays such a critical role in the conceptualization of number, it seemed to be a particularly good medium for studying relations between concept and percept. In particular, size perception was chosen for study because the feature of extent prevails in both visual perception proper and in number conception.

With respect to number, integers are organized in terms of an interval scale with the formal properties of ordinality, equality of intervals and additivity of intervals. By virtue of these properties there is 'numerical distance' between integers, which varies: there is equality of 'numerical distance', e.g. between '1–5' and '5–9'; and there is inequality of 'numerical distance', e.g. between '1–2' and '2–9'. Now, turning to perception *per se*, extent is, of course, a significant property. We have to recognize, however, that there are not only physical factors which define visual extent: two lines, physically equal, may or may not appear equal, depending on a number of subjective, as well as objective, factors. This study assumes that, in

---

* This investigation was supported by research grant MH-00348 from the National Institute of Mental Health, Public Health Service.

particular circumstances, 'number distance' may be a factor which influences the perception of extent.

In order to inquire about the interaction of 'numerical distance' and perceptual extent, a method had to be devised which minimizes the stability of the perceptual experience of extent which occurs under normal circumstances. One way of making extent less stable, i.e. more susceptible to influence, is to present the extent in terms of points which demarcate it successively instead of simultaneously. Accordingly, the method employed consists in making judgements, in a dark room, of visual extents demarcated by successive flashes of light.

Further, the 'number-extent' interaction was studied by having the light flashes appear as number forms, i.e. numbers from 1 to 9. Thus, the observer, in an otherwise dark room, saw three numbers, located along different points in the horizontal dimension, flash one after the other; his task was to judge the visual extent demarcated by the first and second number flash as compared with the visual extent demarcated by the second and third number flash. Given the general assumption that, in these circumstances, there is interdependence between 'numerical distance' and perceptual extent, it is hypothesized that if an extent is demarcated by number, the perception of extent differs depending on the interval between the numbers.

Further, assuming a notion of schema simply as a means for structuring conceptual as well as perceptual events, a specific hypothesis was formulated with the general idea that the magnitude of the interval between the numbers interpenetrates (is assimilative to) the perceptual extent: if a relatively short 'numerical distance' (1–2) is assigned to a particular visual extent it appears relatively shorter than if a relatively long 'numerical distance' (2–9) is assigned to it. In order to make the physical extent to which the numerical interval 1–2 is assigned appear equal to the physical extent demarcated by 2–9, the physical extent demarcated by 1–2 would have to be made bigger.

PROCEDURE

A horizontal extent was demarcated by the numbers 1 and 9. This total extent was divided into two horizontal extents by a third number chosen from the number series 2–8; e.g. in one instance the middle number was 2, so that the left horizontal extent was demarcated by 1 and 2, and the right horizontal extent was demarcated by 2 and 9. In general, the subject's task was to adjust one horizontal extent so that it appeared equal to the other horizontal extent. The middle integer, i.e. the point demarcating the border of the left *vs.* right extent, was always fixed in space in the subject's objective median plane. In some trials the right extent was the standard, and the left extent the variable; in others this was reversed. The subject was required to instruct the experimenter to make the variable extent equal to the standard by moving to left or right the outside integer which demarcated the variable extent.

Three small light boxes, the fronts of which consisted of opaque glass, were used for presenting the stimuli. The stimuli were stencils of integers $2\frac{1}{2}$ in. in height; because the numbers varied in width they were surrounded by a luminous circle $\frac{1}{16}$ in. wide and $2\frac{3}{4}$ in. diameter. The stencil for each integer was mounted on the front of each light box so that the configuration of the integers and a surrounding ring of light was the only thing visible to the subject. These light boxes were arranged in the fronto-parallel plane on a horizontal track 8 ft. from the subject. The central light box was fixed in the objective median plane; the light boxes on the left or right were fixed depending upon whether the standard was on the left or right. Thus two light boxes, either centre and left or centre and right, were fixed; the third light box was moved on the horizontal track by the experimenter upon instruction by the subject.

The integers were flashed for a duration of 0·55 sec. successively from left to right with an interval of 0·45 sec. between the first and second integers and between the second and third integers.

The timing mechanism consisted of a cam propelled by a constant speed motor (20 rev./min.) which activated micro-switches; these in turn controlled lighting of the integers.

The subject was led blindfolded into the experimental room and seated in a specially designed chair with a head rest which fixed his head in the direction of the middle integer. He was then instructed to make the two horizontal extents equal by telling the experimenter how to move the variable integer so that one extent, e.g. 1–2, appeared equal to the other extent, e.g. 2–9. When the integers flashed on the first trial, the subject named the integers from left to right, and then made a judgement concerning the equality of the extents between integers. With a judgement of inequality, the experimenter, upon instructions from the subject, moved one of the integers; the lights were then successively flashed again and the subject made another judgement; this procedure—varying the extent, successive flashing of the three integers, the subject making judgement—was repeated until the two extents appeared equal.

The main condition for perception of extent involved variation of 'numerical distance'. Three sets of integers were selected to demarcate the physical distances, viz. 1–2–9, 1–5–9, and 1–8–9. In any set of three the low number was always on the subject's left and the high number on his right. As one can see, the 'numerical distance' used for the left extent was smaller than, equal to and greater than the 'numerical distance' assigned to the right extent.

Two further conditions were employed: one pertained to the side on which the variable was located; the other to the initial extent of the variable at the start of a trial, i.e. (1) standard on left—variable on right *vs.* standard on the right—variable on the left; and (2) initial extent smaller (25 cm.) than standard (35 cm.) and initial extent larger (45 cm.) than standard (35 cm.). Each subject was tested under all twelve conditions. The distance of the variable from the middle integer served as the measure employed in the analysis.

RESULTS

The results are presented in Table 1 and Fig. 1. In order to evaluate the results of Table 1, the following consideration has to be taken into account: because the integers are always placed in ascending order from left to right, and because the dependent variable involves the 'small numerical distance' on half the trials and the 'large numerical distance' for the other half, the effect of 'numerical distance' is here reflected in the statistical interaction 'Numerical distance × position'. Examination of Table 1 shows a significant statistical interaction of 'numerical distance × position', which means that the effect of 'numerical distance' is significant. The specific effects are presented in terms of mean *apparent extent*, which is defined as the physical distance to which the variable must be adjusted so that it appears equal to the standard.

Examination of the means in Table 1 and the graphic presentation in Fig. 1 shows: when the 'numerical distance' of the variable is small compared with the 'numerical distance' of the standard, the apparent extent of the variable is small; when the 'numerical distance' of the variable is equal to that of the standard, the apparent extent of the variable approaches that of the standard; and when the 'numerical distance' for the variable is larger than that of the standard, the apparent extent for the variable is greater.

In sum, 'numerical distance', i.e. magnitude of the numerical interval assigned to a given visual extent, significantly affects the perceived extent: an extent demarcated by greater 'numerical distance' appears equal to an extent demarcated by smaller 'numerical distance' when the former extent is physically larger relative to the latter; or, with greater 'numerical distance' there is greater apparent extent.

A secondary finding may be noted, namely, that starting extent plays a significant role: when the judgements are initiated with the variable initially set at a short (long) extent, the overall final apparent extent is smaller (larger).

Table 1. *Effect of 'numerical distance' upon 'apparent extent'*

F tests

Source of variation	D.F.	Mean square	*F*	*P*
Individuals	23	1230·33	< 1·00*	N.S.
Sequence	11	868·86	< 1·00*	N.S.
Indiv./seq.	12	1561·67		
Order	11	3533·02	1·96	N.S.
Conditions	11	11887·98	6·64	< 0·05
Numerical distance (*N*)	2	2220·05	1·24	N.S.
Position of variable (*P*)	1	51·84	< 1·00	N.S.
Starting extent (*S*)	1	65673·92	36·68	< 0·01
$N \times P$	2	25842·65	14·43	< 0·01
$N \times S$	2	3796·42	2·12	N.S.
$P \times S$	1	513·67	< 1·00	N.S.
$N \times P \times S$	2	405·47	< 1·00	N.S.
Square uniqueness	110	1250·69	< 1·00	N.S.
Error	132	1790·54		
Total	287	1992·53		

* Tested against indiv./seq.; all others tested against error.

Means (mm.)

	Conditions				
No.	Numerals	Position of variable	Starting extent	Mean extent	Mean extent pooled for† starting extent
1	1–2–9	Left	Small	313·8	325·3
2	1–2–9	Left	Large	336·6	
3	1–2–9	Right	Small	344·4	360·4
4	1–2–9	Right	Large	376·4	
5	1–5–9	Left	Small	345·9	356·6
6	1–5–9	Left	Large	367·3	
7	1–5–9	Right	Small	339·4	347·9
8	1–5–9	Right	Large	356·6	
9	1–8–9	Left	Small	344·4	363·6
10	1–8–9	Left	Large	382·8	
11	1–8–9	Right	Small	309·8	334·8
12	1–8–9	Right	Large	359·3	

† A difference of 17·4 is significant at the 0·05 level.

DISCUSSION

It will be recalled that there were two hypotheses: first, the general hypothesis that if an extent is demarcated by two integers, the perception of extent varies depending on the interval between the two integers employed; secondly, the specific hypothesis that the magnitude of the interval interpenetrates (is assimilative to) the perceptual extent, so that with a relatively short numerical interval (e.g. 1–2) the visual extent appears relatively shorter than with a relatively long numerical interval (e.g. 2–9).

The findings conform with the general hypothesis. However, the results are opposite to those expected from the specific hypothesis. According to this hypothesis, when equating two distances, e.g. one demarcated by 1–2 (short numerical interval) and the other by 2–9 (long numerical interval) the first line (which is hypothesized to appear shorter) would have to be made longer. Our findings are the opposite, i.e. the aforementioned two distances *appear equal* when 1–2 is made *shorter* than 2–9. How can one account for the contradiction between hypothesis and findings?

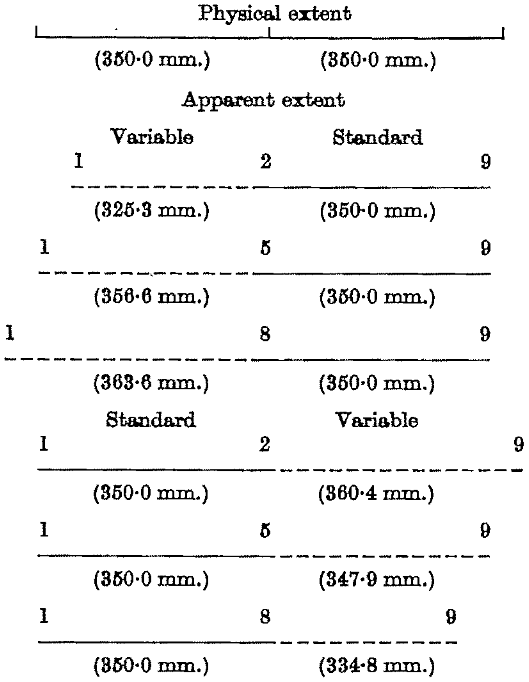


Fig. 1. Schematic representation of results: mean apparent extent.

The resolution of the contradiction hinges on how one views the interaction of cognitively connoted intervals in relation to visual length. The specific hypothesis formulated at the beginning of the study was based on the assumption that a cognitive interval, such as the numerical interval 1–2 or 2–9, has to be transformed into visual distances in order to be experienced. This would then lead to an interaction between the visual distances as given by optic stimuli and the visually transformed numerical interval. Essentially, this means that an interaction between two visual experiences takes place. That is, a visually transformed numerical interval (e.g. 1–2) would influence the visual appearance of a physical length demarcated by 1–2 and, accordingly, this physical length would appear smaller than the physical length which is assimilated to the visually transformed numerical interval 2–9. In sum, there would be assimilation of the perception of visual length to the visual projection of the 'numerical interval'. Since the task is to adjust the line segments 1–2 and 2–9 to equality, this assimilative process would lead to the expectation of the original specific hypothesis. The facts are contradictory to this hypothesis. Moreover, it seems unlikely that in this process of discrimination a visual-visual interaction takes place which is mediated by a conceptual transformation of a numerical interval into a visual schema.

The alternative explanation, which seems feasible to us, hinges on the following analysis of the relationship between a cognitively apprehended numerical interval and its interaction with a sensory event.

First of all, the task involves *visual comparisons* between two extents.

Secondly, a cognitively apprehended numerical interval is not a sensory—here a visual—event. It is an experience which we believe cannot be described merely in sensorial terms; we have to find ways to describe the phenomenal dimensions in terms which have dynamic, vectorial characteristics. Such a vectorial dimension pertinent to the property of extent, we believe, is referred to by the polarity, *expansiveness-contractiveness*.*

Thirdly, we assume that wherever a relationship between 'conceptual' and 'sensorial' phenomena occur, this relationship has to be defined as a 'correspondence'.† We would then maintain that a correspondence is possible between numerical intervals and visual length of lines. In our situation such a correspondence is forced by the assigning of numbers to demarcating visual extents.

Fourthly, the task demands the judgement, in terms of apparent equality, of two lines. This 'apparent equality' can only be defined in terms of their equivalence; the task requires that two such lines should appear equivalent in regard to the numerical interval assigned to them. In our terms, lines would appear equal if their visual length *corresponds* to the numerical interval assigned to them.

Therefore, in order to achieve visual equivalence between the physical length of line assigned 1–2, and the physical length of line assigned 2–9, 1–2 must be made shorter than 2–9. When such an adjustment is made, the relationship between the two physical extents is in *correspondence* with the relationship of the cognitively apprehended numerical intervals.

Such a *correspondence* between a conceptual numerical interval and visual (physically determined) length is attainable because both, from the viewpoint of sensory-tonic theory, have the characteristic of expansiveness: a high (low) degree of expansiveness of the visual line corresponds to a high (low) degree of expansiveness of a numerical interval.

REFERENCES

Bartlett, F. C. (1950). *Remembering. A Study in Experimental and Social Psychology.* Cambridge University Press.

Head, H. (1926). *Aphasia and Kindred Disorders of Speech.* Cambridge University Press.

Piaget, J. (1952). *The Origins of Intelligence in Children* (trans. by Margaret Cook). New York: International Universities Press.

Vernon, M. D. (1955). The functions of schemata in perceiving. *Psychol. Rev.* **62**, 180–92.

---

* This polarity 'expansiveness-contractiveness' has been introduced by us in theoretically interpreting perceived extent in sensory-tonic terms; we assume that this polarity is a reflexion of variation in organismic vectorial (sensory-tonic) state pertinent to the property of extensionality (Werner & Wapner, 1949, 1952, 1956; Wapner & Werner, 1957; Wapner, McFarland & Werner, 1962).

† An everyday illustration of such a correspondence is that between the concept of temperature and the length of the mercury column in a thermometer. Here, the correspondence is based on the common experience that heat can be experienced as a gradual intensification corresponding with 'extensification' on a thermometric scale.

WAPNER, S., MCFARLAND, J. H. & WERNER, H. (1962). Effect of visual spatial context on perception of one's own body. *Brit. J. Psychol.* **53,** 222–30.

WAPNER, S. & WERNER, H. (1957). *Perceptual Development.* Worcester: Clark University Press.

WERNER, H., *Comparative Psychology of Mental Development.* New York: Harper, 1940 2nd ed., Chicago: Follett, 1948; 3rd ed., New York: International Universities Press, 1957; 4th ed., New York: Science Editions, 1961.

WERNER, H. & WAPNER, S. (1949). Sensory-tonic field theory of perception. *J. Personality,* **18,** 88–107.

WERNER, H. & WAPNER, S. (1952). Toward a general theory of perception. *Psychol. Rev.* **59,** 324–38.

WERNER, H. & WAPNER, S. (1956). Sensory-tonic field theory of perception: basic concepts and experiments. *Revista di psicologia,* **50,** 315–37.

(*Manuscript received* 12 *November* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 429–432
Printed in Great Britain

INFORMATION, ACOUSTIC CONFUSION AND MEMORY SPAN

By R. CONRAD and A. J. HULL

Medical Research Council, Applied Psychology Research Unit, Cambridge

Immediately after visual presentation, subjects were required to recall 6-letter sequences. Sequences were drawn from four vocabularies. There were two 3-letter vocabularies, distinguished by the probability of acoustic confusion within them, and two 9-letter vocabularies similarly distinguished. Memory span is shown to be effectively independent of information per item, and to depend substantially on the probability of acoustic confusion within vocabularies.

A recent paper by the author (Conrad, 1964) demonstrated a highly significant association between errors in immediate recall of 6-letter sequences, and listening errors. Letters of the alphabet which are difficult to hear against a white noise background are also difficult to recall, even when presentation for recall is visual. The effect was also shown to hold true for words (Conrad, 1963).

It was argued that, if letters which are acoustically confusing are more likely to be confused in memory, then sequences drawn from acoustically homogeneous vocabularies would be more difficult to recall than those drawn from vocabularies which were acoustically heterogeneous. It was suggested that the acoustic nature of the vocabulary might in fact be a more important variable in determining memory span than the size of vocabulary from which sequences were drawn. The present experiment tests these two hypotheses.

METHOD

Thirty-six 6-letter sequences were constructed from each of four consonant 'vocabularies'. The vocabularies were:

1. J K N
2. C D F H L N Q Y Z
3. F S X
4. B C D G M N P T V

The sequences were based on tables of random numbers with the following constraints:

(*a*) Within each set of thirty-six sequences, letters occurred equally often in each serial position.

(*b*) No letter occurred twice or more in succession within a sequence.

(*c*) The sequential structure of the sequences drawn from vocabularies 1 and 3 and from vocabularies 2 and 4 were identical. For example, if the first sequence from vocabulary 1 was JKNJNK, the first sequence from vocabulary 3 was FSXFXS.

The vocabularies were chosen on the basis of two variables. First, vocabulary size, three or nine letters. Secondly, acoustic confusability within the vocabulary, so that for each size of vocabulary there was one which had high, and one which had low, confusability. This variable requires explanation.

Conrad (1964) presents a confusion matrix for letters of the alphabet spoken against a white noise background. For subsets of this matrix composed of the letters of the above four vocabularies, the probability was determined with which, in a listening task, each letter would be confused with each other in the set. On the assumption that the constant ratio rule (Clarke, 1957) is valid, the mean of these probabilities gives a value which can be used as an index of acoustic confusability within each vocabulary. For the four vocabularies this value was: (1) 0·24, (2) 0·29, (3) 0·51, (4) 0·69.

Sequences were presented visually letter by letter at a rate of eighty letters per minute by means of 16 mm. frame by frame projection. At the end of each sequence the projector was stopped and subjects wrote down what they could remember of the sequence in correct order. The vocabulary from which the sequences were drawn was printed on each answer sheet.

In fact two experiments were carried out using the same test material. Since the results were identical, the experiments will be described together. The first experiment used a factorial design with twenty-four naval ratings as subjects. Subjects were individually tested, and were given one set of thirty-six sequences on each of 4 successive days. This procedure involved the possibility of confusion between vocabularies, particularly since they were not mutually exclusive.

The second experiment used a matched groups design, and subjects were trainee telephonists. Four groups were matched on the basis of age and score on an immediate memory test for 7-digit sequences using the same presentation procedure as in the main test. Because subjects were available only once, matching was achieved merely by increasing the size of each group until the groups no longer differed in terms of mean age (27 years) and digit memory errors (15·7 %). To reach this criterion the necessary group sizes ranged from seventy to eighty subjects per group. In this experiment the subjects of each group memorized sequences drawn from only one of the four vocabularies.

Recorded sequences were scored in terms of letters wrong. A letter was scored wrong if it was wrong for that serial position. Thus a paired transposition counted as two errors; a whole sequence omitted counted as six errors.

RESULTS

The essential results are shown in Fig. 1. It is clear that there is no real difference in the results of the two experiments. On the other hand, since the non-linearity of the relationship between errors and probability of acoustic confusion is evident and similar in both experiments, it must be considered a genuine effect and will be discussed later.

In both experiments, the means of the two 3-letter and of the two 9-letter vocabularies are significantly different at better than the 0·001 level. So that, with vocabulary size (information per item) constant, memory span is affected by the probability of acoustic confusion within the vocabulary set. Indeed, vocabulary size might seem to be irrelevant, since vocabulary 2 is significantly better than vocabulary 3 ($P < 0{\cdot}001$).

Although the differences between vocabularies 1–2 and 3–4 are not significant ($P > 0{\cdot}05$), one would hesitate on the present data alone to claim a simple relationship between span and confusability. There are inevitably structural differences in the 3- and the 9-letter vocabularies. In particular more letters are repeated in the sequences drawn from the smaller set.

There is little formal evidence to go on about the effect of repeated items in a sequence. In one sense repetition could make sequences easier to recall, since a repeated item could aid coding. On the other hand, for this aid to be maximal one would imagine that the repeated items would need to occur in fixed positions in the sequence. This was not of course the case in the present experiment. An attempt to examine this point directly was made by questioning whether subjects were better able to memorize sequences with repeats than those without.

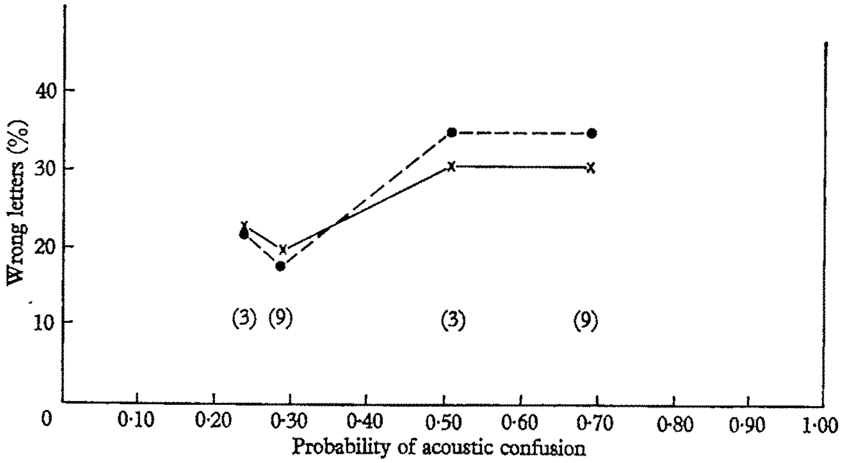


Fig. 1. Immediate recall of 6-letter sequences. × — ×, Matched groups, $N = 70$–80; ● – – ●, Factorial design, $N = 24$.

The 3-letter vocabulary sequences always contained repeated letters. Only the 9-letter vocabulary sequences taken from the matched group experiment were therefore considered. Sequences contained 0, 1 or 2 repeated letters. The mean number of wrong letters per sequence per subject was calculated for these three repetition classes. For neither vocabulary were the differences significant. Thus when the structure (repeated letters) changes from sequence to sequence, subjects apparently ignore repetition information; or at any rate it has no effect on their performance. It is therefore unlikely that structural differences between 3- and 9-letter vocabulary sequences influence the results.

Discussion

Clearly the two hypotheses outlined in the introduction are fully confirmed. The two 3-letter vocabularies and the two 9-letter vocabularies give markedly different spans. Furthermore, comparison of vocabularies 2 and 3 shows a significant difference which is in the wrong direction for any informational prediction, but the right direction from the standpoint of the present argument.

But of course a continuous linear relationship would have been preferred, and there are two sources of sampling error which may be relevant. The probability of acoustic confusion was calculated from a set of empirical data. The relative quantities in these data would depend to some extent, as Miller & Nicely (1955) have shown, on the signal–noise ratio used in the listening test. Thus the abscissa values in Fig. 1 could vary if different listening data were used; though not so much that the relative order of the points would be likely to change.

A second source of error lies in the way the matrix of listening errors (Conrad, 1964) was used; in effect only a single vocabulary out of a number of possible vocabularies of the same size, all having approximately the same probability of acoustic confusion. The results may therefore involve two sampling errors which, built into the sequences used, are independent of subjects or experimental design.

These results are completely at variance with those of Pollack (1953), who reported that 'amount of information transmitted...increases almost linearly as the amount of information per item...is increased'. Pollack's procedures differed in a number of ways from those used here, but the critical difference is that he probably inadvertently increased the acoustic confusability of his vocabularies as he increased their size. This point was discussed more fully by Conrad (1964) in support of the predictions which led to the present experiment, and need not be repeated here.

The new implications of this experiment are largely practical. Whenever subsets of letters or digits, or indeed any other material which can be verbalized, are compared in memory experiments (e.g. Mackworth, 1963) care must be taken to equate the subsets for acoustic confusability. The results otherwise will certainly be contaminated by this factor.

We are indebted to the Post Office and to the Union of Post Office Workers for facilitating this experiment.

REFERENCES

Conrad, R. (1963). Acoustic confusions and memory span for words. *Nature, Lond.*, **197**, 1029–30.

Conrad, R. (1964). Acoustic confusions in immediate memory. *Brit. J. Psychol.* **55**, 75–84.

Mackworth, J. F. (1963). The relation between the visual image and post-perceptual immediate memory. *J. Verbal Learning and Verbal Behaviour*, **2**, 75–85.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Amer.* **27**, 338–52.

Pollack, I. (1953). Assimilation of sequentially encoded information. *Amer. J. Psychol.* **66**, 421–35.

Clarke, F. R. (1957). Constant-ratio rule for confusion matrices in speech communication. *J. Acoust. Soc. Amer.* **29**, 715–20.

(*Manuscript received* 11 *November* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 433–437
Printed in Great Britain

TRANSFER OF TRAINING BETWEEN SPACE-ORIENTED AND BODY-ORIENTED CONTROL SITUATIONS

BY M. HAMMERTON AND A. H. TICKNER
Medical Research Council, Applied Psychology Research Unit, Cambridge

An experiment was designed to determine whether skill in a control task is learned as a pattern of body-oriented limb-movements, or as a system of space-oriented relationships. A control task was set up which could be correctly oriented: A, both bodily and spatially; B, bodily but not spatially; C, spatially but not bodily. Four groups of six subjects were used severally to study transfer of training from A to B, B to A, A to C, and C to A. It was found that transfer to condition C was significantly the worst. It is tentatively concluded that this type of skill is learned mainly as a body-oriented pattern of limb-movements.

I. INTRODUCTION

In training personnel for skilled tasks, it is important to be clear about what exactly is learned. This is particularly important if the practical situation may differ from the training situation. In a review paper, Campbell (1954) concluded that '...in...transposition experiments the results [are] overwhelmingly on the side of interpreting the learned response in habit formation in terms of organism-environment shifts, rather than in terms of limb or body movements in terms of body parameters'. In other words: learned skills are space- or display-oriented, rather than body- or limb-oriented.

In a recent general paper, Gleitman (1963) avoided so definite a pronouncement, and it would seem that the question is still an open one. Opportunities for testing it with human subjects are offered by visual–motor control experiments. There have been a number of studies of skilled control behaviour with varying operator positions and control–display alignments (e.g. Humphries, 1958; Shephard & Cook, 1959; Abbey, 1961, 1962). In these, however, transfer between situations was not a principal object of study. Also, as the experiments were conducted, spatial and body orientations were altered together. The purpose of the present experiment was to alter the space-orientations and body-orientations of the situations separately, and to examine the transfer of training between the several conditions. This was possible because only thumb-movements were made in control, and the display was on a cathode-ray tube (C.R.T.). In these conditions the body frame of reference can reasonably be equated with subject's forearm, and the space frame of reference with the display.

II. METHOD

Subjects. Four groups, each of six Royal Naval Ratings, whose ages ranged from 18 to 26 years, acted as subjects.

Apparatus and task. The apparatus used was the Applied Psychology Research Unit control simulator. The subject sat facing a C.R.T., upon which a spot and a short vertical line of light could be made to appear. The vertical line was the 'target'; the movements of the spot were controlled by the subject. For this control, the subject used a thumb joystick, mounted upon a pistol grip, which he held in his right hand. The control was so mounted that the plane of movement of the subject's thumb was parallel to the plane of the face of the C.R.T.—45° to the horizontal in both cases. First order—i.e. velocity—control was used. The control sensitivity, i.e. spot velocity (radians/sec.)/angular movement of stick (radians), was 0·216 sec.$^{-1}$. The maximum spot velocity was 0·212 rad./sec. The task was a simple acquisition one. When the experimenter pressed a button, the target-line moved to a fixed position in the upper left-hand quadrant of the screen, 3 cm. from the initial position of the line and spot. Simultaneously, the subject's control became operative, and a decatron timer started. The subject had to move his spot as rapidly as possible on to the mid-point of the target line, and to keep it there for 2 sec. This completed the run; whereupon the decatron stopped; spot and line jumped back to their initial position; and the control became inoperative. The reading of the decatron, minus the 2 sec. 'holding time', constituted the subject's score: the shorter the better.

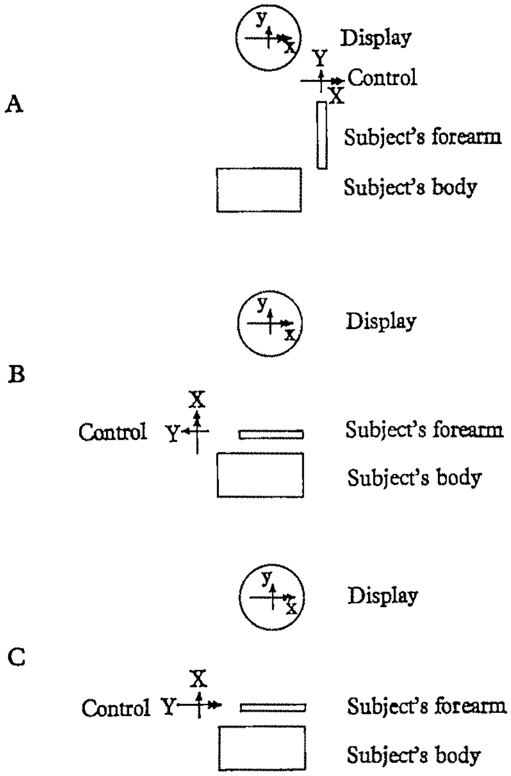


Fig. 1. Schematic layout of experimental conditions. A, control correctly oriented to body (forearm) and to space (display); B, control correctly oriented to body but not to space; C, control correctly oriented to space but not to body.

The pistol grip could be clamped in one of two positions: (i) so that the subject's forearm was pointing straight forward, (ii) so that the subject's forearm was pointing to his left, across his body, at approximately 90° to its direction in the first position.

Let us denote movements of the spot on the screen as in x and y as usual (see Fig. 1). Let us denote side-to-side movements of the thumb (and hence of the stick), relative to the forearm as in X, and fore-and-aft movements of the thumb relative to the forearm as in Y. In the present experiment, three conditions were employed. These were:

Condition A. Stick in position (i); movements in X and Y produced corresponding movements in x and y. Thus the control was correctly oriented both bodily and spatially. (That this is the highly compatible situation is well-known—see, for example, Shackel & Whitfield, 1963. It is, however, easy to be confused by thinking of the controls of an aircraft. An aircraft is, of course, controlled from the inside; the joystick is moved towards the desired inclination of the machine's z-axis, and if the horizon is regarded as the pilot's display, it moves up as the stick is moved forward.)

Condition B. Stick in position (ii). Again movements in X and Y produced corresponding movements in x and y. Thus, the control was correctly oriented bodily, but wrongly oriented spatially.

Condition C. Stick in position (ii). Positive movements in X produced positive movements in y; negative movements in Y produced positive movements in x.

Thus, the control was wrongly oriented bodily, but correctly oriented spatially—i.e. stick movements were parallel to spot vectors. The salient features of the three conditions are shown in Fig. 1.

Procedure. Transfer was studied in four cases: (1) condition A to condition B, (2) B to A, (3) A to C, (4) C to A.

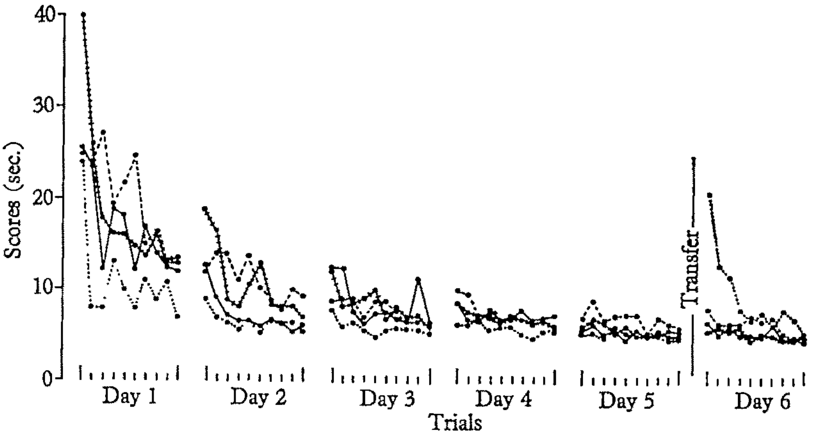


Fig. 2. Curves showing variations of group mean performance during the experiment. Group 1 (●—●) transferred from condition A to condition B, group 2 (●+++●) from B to A, group 3 (●...●) from A to C, group 4 (●---●) from C to A.

A separate group of subjects was used for each of these. In each case, subjects had 10 trials per day for 5 days in the training condition, followed, on the 6th day, by 10 trials in the transfer condition. In order to eliminate effects due to 'loaded' instructions, these were always given to the subjects in such terms as emphasized the 'naturalness' of the condition. Thus, in condition A, they were told: 'To move the spot from side to side, you move your thumb from side to side; to move it up and down you move your thumb forward and back.' The same form of words was used for condition B. In condition C, they were told: 'You move your thumb in the same direction that you want the spot to move. If you move your thumb towards the wall, the spot will move towards the wall, and so on.' In all cases, the task was demonstrated once before trials began.

To ensure that the subjects had grasped what was required, they were asked to explain the control to the experimenter. None found any difficulty.

III. RESULTS

The mean performance of the several groups during the course of the experiment is shown in Fig. 2. Each point on the curves is thus the mean of six subjects' scores.

For each group, the measure of transfer used was ρ, defined as

$$\rho = \frac{F-T}{F-L},$$

where F is the performance on the first training trial, L on the last, and T is the performance on the first transfer trial. Group mean ρ values, and variances, are shown in Table 1. It is evident upon inspection of the curves, and of the table, that it was very much more difficult to transfer to condition C than to conditions A and B. Analysis of variance showed that the ρ values obtained by group 3 differed significantly from those of the other groups ($P < 0{\cdot}001$). This was confirmed by t-tests between all six possible pairs of groups. The ρ values of groups 1, 2 and 4 belong, therefore, to the same population; those of group 3 do not.

It is also noticeable that on the initial trial of the first 2 days, the performance of group 2 (B–A) seems worse than that of the other groups. This difference was examined, using the F-test, and was found to be significant ($P < 0{\cdot}05$).

Table 1. *Means and variances of transfer measures* (ρ) *for experimental groups*

Group	Mean ρ	Variance of ρ
1 (A–B)	0·90	0·010
2 (B–A)	0·99	0·001
3 (A–C)	0·05	0·265
4 (C–A)	0·86	0·016

IV. DISCUSSION

As was remarked in the introduction, when using a thumb-joystick of the kind described here, the body frame of reference may reasonably be equated with the forearm. Similarly, the space frame of reference is equivalent to the display. Thus we have examined transpositions between situations which were correctly oriented to space and body, and situations correctly oriented to only one of these.

The results show very forcibly that, in the present context, it is incorrect body orientation that produces a marked decrement of performance. This, of course, bears strongly against Campbell's conclusion quoted at the beginning; and supports the contention that 'what is learned' in skilled behaviour is a pattern of limb movements defined in terms of body parameters. The good transfer from condition C to condition A is presumably an instance of the familiar situation of transferring from a more to a less difficult situation.

The poor first trial performances on days 1 and 2 by group 2, are open to two possible explanations. It might be that there was an initial expectation, or population stereotype, of space-orientation in such controls, possibly acquired from previous experience with related skills. Such stereotypes are often regarded as important when operators have to use, perhaps in an emergency, controls on which they have had little practice. However, the present findings clearly show that, within the limits of the tasks used, such expectations are very rapidly overcome.

On the other hand, it is possible—and the authors incline to this view—that the significance was a spurious one. Early in training, performances are fluctuating widely, and it would not be surprising if, on isolated trials, a particular group should produce scores which differ sharply from those of the others. For example, on the 2nd and 3rd trials of day 1, group 3 is markedly better than the others, the F-test again showing significance ($P < 0{\cdot}05$), although one of the other groups used the same control.

On a practical plane, it seems that if it is necessary to have a control that is not doubly compatible (as in the desirable condition A), then it is more important to maintain body-orientation than space-orientation. This situation may well arise, e.g. in the design of space vehicles. Here the pilot may have to adopt a posture which compels the use of thumb-joysticks mounted at arbitrary orientations to the movements of the vehicle. The frame of reference which he can most readily be trained to use, then, would seem to be his own forearm.

REFERENCES

ABBEY, D. S. (1961). Degree of C/D alignment and performance on a complex perceptual-motor task. *Percep. Motor Skills*, **13**, 343–8.

ABBEY, D. S. (1962). Partial reversal of C/D relations, body position, and performance on a complex perceptual-motor task. *Percep. Motor Skills*, **14**, 34.

CAMPBELL, D. T. (1954). Operational definition of 'what is learned' via the transposition experiment. *Psychol. Rev.* **61**, 167–74

GLEITMAN, H. (1963). Place learning. *Sci. Amer.* **204**, 116–22.

HUMPHRIES, M. (1958). Performance function of control-display relations, positions of the operator, and locations of the control. *J. appl. Psychol.* **42**, 311–16.

SHACKEL, B. & WHITFIELD, D. (1963). *Instruments and People*. London: H.M.S.O.

SHEPHARD, A. H. & COOK, T. W. (1959). Body orientation and perceptual-motor performance. *Percep. Motor Skills*, **9**, 271–80.

(*Manuscript received* 2 *December* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 439–446
Printed in Great Britain

EXPERIMENTAL STUDIES IN 'HYPNOTIC' BEHAVIOUR: SUGGESTED DEAFNESS EVALUATED BY DELAYED AUDITORY FEEDBACK*

BY T. X. BARBER AND D. S. CALVERLEY
Medfield Foundation, Harding, Massachusetts, U.S.A.

Delayed auditory feedback (DAF) was used to assess the objective consequences of suggested deafness. Forty-two women students were randomly assigned to three groups of fourteen subjects. Group A received suggestions of deafness subsequent to a standardized 15 min. hypnotic induction procedure; Group B received suggestions of deafness under waking conditions; Group C served as a control. Suggestions of deafness, given with *or without* a preceding hypnotic induction procedure, were effective with the majority of subjects in eliciting subjective reports of altered auditory perception. However, subjects claiming partial or total 'subjective deafness' responded to the DAF with stuttering, mispronunciations, increased vocal intensity and slowed rate of speaking in the same way as the controls who heard normally.

General texts on hypnotism assert or imply that suggestions of deafness are often effective with hypnotized individuals but are rarely if ever effective with persons who are not hypnotized. This assertion appears to derive from unsystematic observations rather than from carefully controlled experimental studies. Although Erickson (1938), Kline, Guze & Haggerty (1954), Malmo, Boag & Raginsky (1954), Pattie (1950), and Sutcliffe (1961) administered suggestions of deafness to selected hypnotic subjects, no one, to our knowledge, has ever given similar suggestions to a non-hypnotized group. The present study is designed to fill this lacuna and to answer the following questions:

(1) Are suggestions of deafness more effective when given under hypnotic conditions than under non-hypnotic conditions in producing total or partial 'subjective deafness', that is, in eliciting testimony that auditory stimuli were either not perceived or were perceived as reduced in intensity?

(2) Does 'subjective deafness' produce *objective* consequences which are similar to or identical with the objective effects of actual deafness?

In the present experiment the objective effects of suggested deafness are evaluated by the method of delayed auditory feedback (DAF) in which the subject's ongoing speech is recorded, delayed for a fraction of a second, amplified, and returned to his ears through a headphone (Chase, Sutton & First, 1959; Smith, 1962; Yates, 1963). Previous studies demonstrated that: (i) The speech of a person with normal hearing is markedly impaired when he hears himself speak within 0·1–0·8 sec. after uttering. This impairment typically includes: speech disturbances such as stuttering and mispronunciations; increased vocal intensity; and decreased rate of speaking. (ii) Totally deaf individuals do not show speech impairment with DAF (Gibbons & Winchester, 1957). (iii) When an individual who claims deafness manifests speech disturbances, increased intensity of utterance, and decrease in rate of speech with DAF it is highly probable that he is malingering (Hanley & Tiffany, 1954).

---

* This research was supported by grants (MY 4825 and MH 7003) from the National Institute of Mental Health, United States Public Health Service.

METHOD

Test materials

All subjects were required to read aloud three paragraphs from Form 1 and three equivalent paragraphs from alternate Form 2 of the Gilmore (1951) Oral Reading Test with DAF. With respect to Form 1, Paragraph I contained 26 words of very easy reading material (standardized for first-grade students), Paragraph VI included 107 words of medium difficulty (standardized for sixth-grade students), and Paragraph X included 252 words (standardized for tenth-grade students). Paragraphs I, VI and X on Form 2 were very similar in number of words, per cent of polysyllabic words, and grammatical construction.

The DAF was produced by means of a modified Bell & Howell magnetic tape recorder with amplification apparatus (Lafayette Instrument Co., Model no. 775A). The subject wore tight-fitting, rubber-cushioned earphones and spoke into a microphone fixed 12 in. from his chin. The output from the microphone was delayed by 0·4 sec. and then fed into an amplification system which added 61–74·5 db. to the speech signal before it was passed into the earphones.

Subjects

The subjects were forty-two female undergraduates from two colleges. Half were volunteers (paid two dollars) and the others participated in the experiment without remuneration to fulfil the requirements of a course in elementary psychology. None had participated in our previous experiments.

When recruited for the experiment, the subjects were told that one-third, selected at random, would be hypnotized and the others would be allocated to control groups and would not be hypnotized; and that, to conduct the experiment properly, it was necessary that subjects should not discuss it with each other. The latter admonition was also repeated to each subject individually at the termination of her own experimental session.

The subjects were randomly assigned to Groups A, B, and C with fourteen to each group. Each subject was tested individually by one experimenter (D. S. C.). All experimental sessions were held in one room provided by each of the colleges.

Procedure

The procedures for Groups A, B and C were as follows:

Group A, hypnotic induction procedure plus suggestions of deafness. Subjects allocated to Group A were first told that they were to be hypnotized. They were then asked to fixate on a light blinking in synchrony with the sound of a metronome, and a standardized 15 min. hypnotic induction procedure, patterned after the hypnosis procedures of Friedlander & Sarbin (1938), Marcuse (1959, pp. 52–3), and Weitzenhoffer & Hilgard (1959), was administered. At the completion of the hypnotic procedure each subject manifested one or more characteristics—e.g. lack of spontaneity and initiative, disinclination to talk—which, according to the consensus of authorities (Erickson, Hershman & Secter, 1961, pp. 55–8; Pattie, 1956, p. 21; Weitzenhoffer, 1957, pp. 210–12) presumably indicate that a person is in 'hypnotic trance'.

Upon completion of the induction procedure, the subject was told that she would remain in deep hypnosis as she performed forthcoming tasks. She was then asked to open her eyes, instructed to read three paragraphs from Form 1 of the Gilmore Oral Reading Test, and told to continue the reading even though she would hear her voice coming back through the earphones as she spoke. The earphones were appropriately placed, and the subject read Form 1 with DAF. The earphones were then removed, the subject was asked to close her eyes and, for a period of 1 min., was given additional suggestions of relaxation, drowsiness, sleep, and deep hypnosis. Suggestions of deafness for the auditory feedback were then given as follows:

> "In a moment I am going to ask you to put the earphones on again and to read three additional passages. When you put the earphones on this time you will be deaf to your voice, you will not hear your voice coming back to you, you will not hear anything at all. As long as the earphones remain on you will not hear your voice, you will be deaf, totally deaf, unable to hear anything at all. When I take the earphones off, you will hear normally again. Now I will put the earphones on and I will point to passages that I want you to read. As long as the

> earphones remain on you will not hear your voice, you will hear nothing at all, you will be deaf, totally deaf."

The earphones were replaced and the subject read the three paragraphs from Form 2 of the Gilmore Oral Reading Test with DAF. After reading Form 2, the subject was given suggestions of relaxation, drowsiness and sleep for 1 min., and was then assessed on response to the Barber Suggestibility Scale (Barber & Calverley, 1963*a*; Barber & Glass, 1962), which includes eight standardized test-suggestions (e.g. suggestions of body immobility, inability to speak and selective amnesia).

Group B, waking suggestions of deafness. Subjects allocated to Group B were first told that they had been assigned to a control group and would not be hypnotized. They were then asked to read the three paragraphs from Form 1 of the Gilmore Oral Reading Test and were told to read continuously even though they would hear their voice coming back through the earphones as they spoke. When the reading of Form 1 with DAF was completed, the earphones were removed and the subject was given suggestions of deafness for the auditory feedback, as follows:

> "I am going to ask you to read three additional passages with the earphones on. This time I want you to ignore and be deaf for your voice coming back to you. Everyone can read better and more fluently a second time by trying very hard during every second to ignore and be deaf for their voice coming back to them. You can do much better this time but you must try hard, try very hard. It can be done, so please, I do not want you to disappoint me. Keep your thoughts directly on the reading material and do not allow yourself to hear your voice in the earphones".

The earphones were replaced and the subject read the three paragraphs from Form 2 with DAF. After this, the subject was given instructions for a period of 1 min. designed to produce positive motivation to perform well on further tasks, as described verbatim by Barber & Calverley (1962, 1963*b*), and the Barber Suggestibility Scale was administered and scored.

Group C, control. Subjects allocated to Group C were first told they had been assigned to a control group and would not be hypnotized. They were next told to read the three paragraphs from Form 1 of the Gilmore Test and to continue reading even though they would hear their voice in the earphones. One minute after reading Form 1 with DAF, they were asked to read Form 2, also with DAF.

Assessment of 'Subjective deafness'

Immediately after completing the above, each subject was given the following rating scale designed to appraise subjective response to suggestions of deafness ('subjective deafness' scale):

> Check one of the following in reference to the second set of readings: (A) I did not hear anything coming through the earphones when I was asked to read a second time; (B) I heard some sounds coming through the earphones during the second reading, but the sounds were muffled, or far away, or not as loud or intense during the second reading as they were during the first reading; (C) I heard my voice coming through the earphones as loud and intense during the second reading as during the first.

'Subjective Deafness' scores of 3, 2, and 1 were assigned to items A, B and C, respectively.

Assessment of objective effects of 'subjective deafness'

To determine whether 'subjective deafness' minimizes or reduces the effects on speech which are normally produced by DAF, scores were assigned to the following variables: number of speech disturbances; intensity of voice; time required to read each paragraph. The speech disturbances, intensity, and time scores were derived from the transcribed responses by replay of each tape recording. The speech disturbances scores were computed by counting the number of words in each paragraph that were mispronounced, stuttered, omitted, or repeated. The intensity scores were derived by dividing each paragraph into six equally spaced units; recording the peak intensity during the reading of each unit (in db. above an arbitrary reference point); and then averaging the six peak intensities to obtain one overall average intensity score for each paragraph. The time scores were obtained by determining the number of seconds required to read each paragraph.

To check the reliability of the scoring, the records of twelve subjects with identifying clues removed were rescored by the original scorer after an interval of one week. The product-moment correlations between first scoring and rescoring were 0·76, 0·92, and 0·99 for speech disturbances, intensity, and time scores, respectively.

RESULTS

Effects of hypnotic induction procedure per se

Before reading the three paragraphs on Form 1 of the Gilmore Oral Reading Test, Group A, but not Groups B and C, had received a standardized 15 min. hypnotic induction procedure, but suggestions of deafness had not as yet been administered. To determine whether the hypnotic induction procedure *per se* affected oral reading performance, one-way analyses of variance were performed separately on the speech disturbances, intensity, and time scores on each of the three paragraphs on Form 1, and Duncan (1955) range tests were used to make multiple comparisons among the means of the three groups. These analyses showed that, with one minor exception, subjects receiving the hypnotic induction procedure (Group A) did not differ significantly from subjects not receiving the hypnotic induction (Groups B and C) in oral reading performance with DAF. (The exception was that Group A read Paragraph I more slowly than Groups B and C.)

Table 1. *Number (and percentage) of subjects scoring* 3, 2 *and* 1 *on 'subjective deafness' scale*

	Scores on 'subjective deafness' scale		
	3 ('did not hear anything')	2 ('sounds were muffled')	1 ('heard normally')
Group A (hypnotic induction plus suggestions of deafness)	0 (0 %)	9 (64 %)	5 (36 %)
Group B (waking suggestions of deafness)	2 (14 %)	11 (79 %)	1 (7 %)
Group C (control)	0 (0 %)	1 (7 %)	13 (93 %)

'Subjective deafness'

Before reading Form 2 of the Gilmore Test with DAF, Group A had received a hypnotic induction procedure plus suggestions of deafness, Group B had received suggestions of deafness *without* a hypnotic induction, and Group C had received neither suggestions of deafness nor a hypnotic induction. Table 1 shows that the suggestions of deafness produced a measurable effect in approximately two-thirds of the subjects in Group A and almost all of the subjects in Group B.

Analysis of variance was performed upon the 'subjective deafness' scores and Duncan range tests were used to make comparisons among the three means. The variance analysis showed a highly significant treatment effect ($F = 13{\cdot}08$; D.F. $= 2/39$; $P < 0{\cdot}001$). The Duncan tests showed that the mean 'subjective deafness' score (2·07) of Group B (waking suggestions) was significantly higher ($P < 0{\cdot}05$) than the mean score (1·64) of Group A (hypnotic induction) and was also significantly higher ($P < 0{\cdot}001$) than the mean score (1·07) of Group C (control); and the mean score of Group A was significantly higher ($P < 0{\cdot}01$) than the mean score of Group C.

Relation of suggestibility to 'subjective deafness'

Of the fourteen subjects in Group A (hypnotic induction), three were rated as 'excellent' subjects (passing all 8 test-suggestions on the Barber Suggestibility Scale), five were rated as 'good' subjects (passing 6 or 7 of the test-suggestions), and the remaining six were rated as 'fair' subjects (passing 2 to 6 of the suggestions). Scores on the Barber Suggestibility Scale and on the 'subjective deafness' scale were positively correlated for this group ($r = 0{\cdot}73$, $P < 0{\cdot}01$), with the more suggestible subjects claiming more 'subjective deafness' than the less suggestible ones.

Of the fourteen subjects in Group B (waking suggestions), six were rated as 'excellent', two as 'good', and six as 'fair' by the criteria mentioned above. Subjects in Group B rated higher on the suggestibility scale showed a tendency (non-significant) to obtain higher 'subjective deafness' scores ($r = 0{\cdot}38$, $P < 0{\cdot}20$).

Objective consequences of 'subjective deafness'

As noted above, nine subjects in Group A and thirteen in Group B manifested some degree of 'subjective deafness', stating that, when they were reading Form 2 of the Gilmore Test with DAF, they either did not hear anything or the sounds were perceived as reduced in intensity (see Table 1). Did these subjects show less speech impairment than the thirteen subjects in Group C (control) who stated that they heard the auditory feedback normally? The treatment scores (scores on Form 2) adjusted for initial differences among subjects (scores on Form 1) were treated by analysis of covariance, which showed that the three subgroups of nine, thirteen and thirteen subjects in Groups A, B and C respectively, did not differ significantly from each other in speech disturbances, in vocal intensity, or in time required to read each of the three paragraphs on Form 2 of the Gilmore Test with DAF.

Analysis of covariance was also performed to determine whether the two subjects in Group B who stated that they did not hear anything at all when reading Form 2 with DAF differed in reading performance from the thirteen subjects in Group C who stated that they heard the auditory feedback normally. This analysis showed that the subjects manifesting 'total subjective deafness' and the subjects with normal hearing did not differ significantly in speech disturbances, intensity, or time scores under conditions of DAF.

DISCUSSION

'Subjective deafness'

Suggestions of deafness were effective with the majority of subjects in the hypnotic group (Group A) and in the non-hypnotic group (Group B) in eliciting reports that objectively loud and disturbing sounds were perceived as reduced in intensity. This finding is generally consistent with the results obtained in a series of recent experiments (Barber & Calverley, 1962, 1963*b*, *c*, 1964*a*, *b*; Barber & Hahn, 1962) and in a series of previous studies, reviewed by Barber (1961*a*, *b*, 1962*a–d*, 1963), which indicate that: (1) It is not necessary to administer a procedure of the type traditionally subsumed under the term 'hypnotic induction' to elicit a high level of suggestibility;

and (2) a high level of response to a variety of suggestions—e.g. suggestions of limb and body rigidity, amnesia, dreaming on a specified topic, analgesia to noxious stimulation, time distortion, visual or auditory hallucinations—can be elicited by administering brief task-motivating instructions, i.e. by telling the subject that he can perform well and he is expected to perform to the best of his ability on suggested tasks.

Although, in the present experiment, the waking group (Group B) obtained significantly higher 'subjective deafness' scores than the hypnotic group (Group A), we cannot conclude, from this finding alone, that suggestions of deafness are more effective when given without a preceding hypnotic induction procedure. Groups A and B received somewhat different suggestions of deafness. Subjects in the hypnotic group were told peremptorily that they would be deaf, whereas subjects in the waking group were told permissively: 'I want you to ignore and be deaf for your voice coming back to you....It can be done, so please, I do not want you to disappoint me.' It may be that suggestions of deafness are more effective when given permissively. Further studies are needed to assess the relative effectiveness of permissive and non-permissive suggestions and to determine whether a hypnotic group is as responsive as a waking group to suggestions of deafness when the suggestions are given in an identical manner to both groups by means of a tape-recording.

Although suggestions of deafness were effective with the majority of subjects in eliciting reports that the auditory feedback was perceived as reduced in intensity, they were effective with only a small proportion of subjects (7 %) in eliciting reports that the sounds were not heard at all. None of the fourteen subjects in Group A and only two of the fourteen in Group B claimed 'total subjective deafness'. It is difficult to argue that we failed to elicit 'total subjective deafness' from a larger proportion of subjects in that we did not use sufficiently 'good' or sufficiently 'suggestible' subjects. Nine subjects (three in Group A and six in Group B) were 'excellent' subjects, passing all eight of the test-suggestions on the Barber Suggestibility Scale. Although responding positively to all other test-suggestions, seven of these nine subjects did not completely accept the suggestions of deafness, stating that, although they perceived the sounds as reduced in intensity, they nevertheless perceived the auditory stimulation. These results are generally in line with those reported by previous investigators. Erickson (1938), for example, succeeded in eliciting apparently 'total subjective deafness' in only six individuals selected from an original group of more than 100 'trained' hypnotic subjects.

Objective effects of 'subjective deafness'

Under the conditions of the present experiment, neither partial nor total 'subjective deafness' produced objective effects which resemble those produced by actual deafness. Subjects manifesting 'partial subjective deafness', stating that auditory stimulation was perceived as reduced in intensity, and subjects stating that they heard normally did not differ significantly in number of speech disturbances, in vocal intensity, and in speed of oral reading with DAF. Further, subjects manifesting 'total subjective deafness', stating that they did not hear anything at all, and subjects stating that they heard normally showed about the same degree of stuttering mispronunciations, loudness of voice, and slowed rate of speaking.

These findings are generally consistent with the results of a recent experiment by Sutcliffe (1961) in which suggestions of deafness given to hypnotized subjects did not prevent the impairment in oral reading performance which normally results from DAF. In the Sutcliffe experiment, as in the present experiment, subjects manifesting either partial or total 'subjective deafness' performed as poorly as the control subjects who heard normally.

REFERENCES

BARBER, T. X. (1961*a*). Antisocial and criminal acts induced by 'hypnosis': a review of experimental and clinical findings. *Arch. Gen. Psychiat.* **5**, 301–12.

BARBER, T. X. (1961*b*). Physiological effects of 'hypnosis'. *Psychol. Bull.* **58**, 390–419.

BARBER, T. X. (1962*a*). Experimental controls and the phenomena of 'hypnosis': a critique of hypnotic research methodology. *J. Nerv. Ment. Dis.* **134**, 493–505.

BARBER, T. X. (1962*b*). Hypnotic age regression: a critical review. *Psychosom. Med.* **24**, 286–99.

BARBER, T. X. (1962*c*). Toward a theory of hypnosis: posthypnotic behaviour. *Arch. Gen. Psychiat.* **7**, 321–42.

BARBER, T. X. (1962*d*). Toward a theory of 'hypnotic' behavior: the 'hypnotically induced dream'. *J. Nerv. Ment. Dis.* **135**, 206–21.

BARBER, T. X. (1963). The effects of 'hypnosis' on pain: a critical review of experimental and clinical findings. *Psychosom. Med.* **25**, 303–33.

BARBER, T. X. & CALVERLEY, D. S. (1962). 'Hypnotic behavior' as a function of task motivation. *J. Psychol.* **54**, 363–89.

BARBER, T. X. & CALVERLEY, D. S. (1963*a*). 'Hypnotic-like' suggestibility in children and adults. *J. Abnorm. Soc. Psychol.* **66**, 589–97.

BARBER, T. X. & CALVERLEY, D. S. (1963*b*). The relative effectiveness of task-motivating instructions and trance-induction procedure in the production of 'hypnotic-like' behaviors. *J. Nerv. Ment. Dis.* **137**, 107–16.

BARBER, T. X. & CALVERLEY, D. S. (1963*c*). Toward a theory of hypnotic behavior: effects on suggestibility of task motivating instructions and attitudes toward hypnosis. *J. Abnorm. Soc. Psychol.* **67**, 557–65.

BARBER, T. X. & CALVERLEY, D. S. (1964*a*). An experimental study of 'hypnotic' (auditory and visual) hallucinations. *J. Abnorm. Soc. Psychol.* **68**, 13–20.

BARBER, T. X. & CALVERLEY, D. S. (1964*b*). Toward a theory of 'hypnotic' behavior: an experimental study of 'hypnotic time distortion'. *Arch. Gen. Psychiat.* **10**, 209–16.

BARBER, T. X. & GLASS, L. B. (1962). Significant factors in hypnotic behavior. *J. Abnorm. Soc. Psychol.* **64**, 222–8.

BARBER, T. X. & HAHN, K. W., Jr. (1962). Physiological and subjective responses to pain producing stimulation under hypnotically-suggested and waking-imagined 'analgesia'. *J. Abnorm. Soc. Psychol.* **65**, 411–18.

CHASE, R. A., SUTTON, S. & FIRST, D. (1959). Bibliography: delayed auditory feedback. *J. Speech Hear. Dis.* **24**, 193–200.

DUNCAN, D. B. (1955). Multiple range and multiple *F* tests. *Biometrics*, **11**, 1–42.

ERICKSON, M. H. (1938). A study of clinical and experimental findings on hypnotic deafness. I. Clinical experimentation and findings. *J. Gen. Psychol.* **19**, 127–50.

ERICKSON, M. H., HERSHMAN, S. & SECTER, I. I. (1961). *The Practical Application of Medical and Dental Hypnosis.* New York: Julian Press.

FRIEDLANDER, J. W. & SARBIN, T. R. (1938). The depth of hypnosis. *J. Abnorm. Soc. Psychol.* **33**, 453–75.

GIBBONS, E. W. & WINCHESTER, R. A. (1957). A delayed sidetone test for detecting uniaural functional deafness. *Arch. Otolaryngol.* **66**, 70–8.

GILMORE, J. V. (1951). *Gilmore Oral Reading Test.* New York: Harcourt, Brace and World.

HANLEY, C. N. & TIFFANY, W. R. (1954). Auditory malingering and psychogenic deafness. *Arch. Otolaryngol.* **60**, 197–201.

KLINE, M. V., GUZE, H. & HAGGERTY, A. D. (1954). An experimental study of the nature of hypnotic deafness: effects of delayed speech feed-back. *J. Clin. Exper. Hypnosis*, **2**, 145–56.

MALMO, R. B., BOAG, T. J. & RAGINSKY, B. B. (1954). Electromyographic study of hypnotic deafness. *J. Clin. Exper. Hypnosis*, **2**, 305–17.

MARCUSE, F. L. (1959). *Hypnosis: Fact and Fiction.* Baltimore: Penguin Books.

PATTIE, F. A. (1950). The genuineness of unilateral deafness produced by hypnosis. *Amer. J. Psychol.* **63**, 84–6.

PATTIE, F. A. (1956). Methods of induction, susceptibility of subjects, and criteria of hypnosis. In Dorcus, R. M. (ed.): *Hypnosis and its Therapeutic Applications*, ch. 2. New York: McGraw-Hill.

SMITH, K. U. (1962). *Delayed Sensory Feedback and Behavior.* Philadelphia and London: W. B. Saunders.

SUTCLIFFE, J. P. (1961). 'Credulous' and 'skeptical' views of hypnotic phenomena: experiments on esthesia, hallucination, and delusion. *J. Abnorm. Soc. Psychol.* **62**, 189–200.

WEITZENHOFFER, A. M. (1957). *General Techniques of Hypnotism.* New York: Grune and Stratton.

WEITZENHOFFER, A. M. & HILGARD, E. R. (1959). *Stanford Hypnotic Susceptibility Scale.* Palo Alto: Consulting Psychologists Press.

YATES, A. J. (1963). Delayed auditory feedback. *Psychol. Bull.* **60**, 213–32.

(*Manuscript received* 21 *September* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 447–452
Printed in Great Britain

VALIDATION OF THE MAUDSLEY PERSONALITY INVENTORY E SCALE

By T. M. CAINE
Claybury Hospital, Woodford Green, Essex
AND K. HOPE*
M.R.C. Unit for Research on the Epidemiology of Psychiatric Illness, University of Edinburgh

The Maudsley Personality Inventory (MPI) E scale is validated by a direct method which has not hitherto been employed. It is shown that, in a neurotic sample, it measures a dimension of personality which has been called the hysteroid-obsessoid dimension and is unrelated to type of illness.

INTRODUCTION

Eysenck has reported on the 'construct validity' (Eysenck, 1957) and the 'clinical validity' (Eysenck, 1959) of the E scale of the MPI. The latter type of validation, employing clinical groups thought to fall at different relative points on the introversion-extraversion dimension, has attracted considerable criticism (Foulds, 1961, 1962; Ingham, 1962). Eysenck & Eysenck (1963) have reported a more direct validation employing extreme extraverts and introverts. We report here a direct validation study which differs from this last (*a*) in sampling technique, (*b*) in the type of subject tested, (*c*) in the criterion employed, and (*d*) in the type of statistical analysis applied.

SAMPLE

The MPI was administered to a consecutive series of fifty patients admitted to Claybury Hospital who were diagnosed by the admitting psychiatrist as neurotic. The average age of the patients was 34·2 years. By chance there were equal numbers of men and women. The average score on the E scale was 23·6, and on the HOQ (see below) it was 22·3. Each patient was assigned to one of the subgroups of neurosis by a psychiatrist. Three patients who were given a diagnosis of anxiety-hysteria were included with the hysterics.

Table 1 shows the mean scores of the various categories of neurotics, together with those reported by Eysenck (1962). Patients with a diagnosis of personality disorder are regarded as belonging to the hysteric side of the dysthymic-hysteric dichotomy because psychiatrists use the terms 'psychopathy' and 'personality disorder' almost interchangeably (Eysenck implies that the introverted score obtained by his personality disorders was predicted; he does not give reasons for such a prediction). The most obvious discrepancy between the two samples is the difference between the means of the two psychopathic groups. An explanation of this datum is offered below. It is also of note that there is no significant difference between dysthymics and hysterics in the present sample.

The Hysteroid Obsessoid Personality Questionnaire (HOQ) was also administered to the patients. The HOQ is a test designed to measure eleven different traits which

---

* The second author held a Medical Research Council Assistantship during this investigation.

are mentioned by Lewis & Mapother (1941) as defining the hysteroid-obsessoid dimension of personality. Unlike the MPI the HOQ was not derived from factor analytic studies but is based on clinically observed empirical correlations. Only the total score on the HOQ is used in the present study. This was validated on a sample of neurotics. In the validation study each patient was rated on each of the eleven traits by a number of observers and the total score on the HOQ was correlated with a personality score derived from the ratings. The correlation was +0·68. The correlation between the HOQ and the E scale in the present sample is +0·70. The only reliability data so far available are derived from a small sample of twenty normal subjects who have been retested after an interval of a year. The reliability coefficient is +0·92. A sample of sixty-two neurotics retested after 6 weeks of intensive therapy gave a coefficient of +0·77. Twenty-two neurotics tested at the conclusion of treatment and retested a year later gave a coefficient of +0·76. Various considerations suggest that, for purposes of categorization, a person who scores 24 or more on the HOQ should be allotted to the hysteroid class, those who score less being classified as obsessoids (Caine & Hawkins, 1963).

Table 1. *Comparison of mean E scores of patients in the present sample with those reported by Eysenck* (1962)

	Sample			
	Caine & Hope		Eysenck	
Diagnosis	*n*	Mean	*n*	Mean
Obsessional state	2	22·5	5	17·4
Anxiety state	26	24·1	20	19·0
Neurotic depression	9	21·9	19	21·9
Dysthymia	37	23·5	44	20·1
Hysteria	7	26·3	29	22·2
Psychopathy	2	15·0	10	31·5
Personality disorder	4	24·3	17	19·7
Hysteria	13	23·9	56	23·1

Table 2. *Mean scores on the MPI E scale when patients are grouped according to personality and diagnosis*

	Personality	
Diagnosis	Obsessoid	Hysteroid
Obsessional state	16·0	29·0
Anxiety state	19·5	31·5
Neurotic depression	20·0	25·7
Hysteria	18·7	32·0
Psychopathy	15·0	—
Personality disorder	5·0	30·7

Table 2 shows the group means when the sample is split according to personality type as well as diagnosis. Three considerations suggest that the E scale measures the same dimension as the HOQ: (1) within each diagnostic group the obsessoids score less than the hysteroids; (2) all obsessoid groups score less than all hysteroid groups; (3) both members of the psychopathic group are obsessoids and this explains the discrepancy in Table 1.

VALIDATION

It should be noted that, although the second consideration entails the first, they have been stated separately. This brings out an important methodological point which has not been sufficiently emphasized. The method of criterion groups, as applied to the E scale, has usually involved showing that hysterics obtain a higher score than dysthymics. All that is shown by this finding is that there is a positive between-groups correlation between the diagnostic dichotomy and the scale. It is quite possible to find a high positive between-groups correlation accompanied by a negative or zero within-groups correlation. In such circumstances it is not reasonable to present the findings as proof that the test in question is a measure of a 'dimension', in any full sense of the word.

Table 3. 2×6 *analysis of MPI E scores of fifty neurotic patients classified according to personality and diagnosis*

Source	D.F.	Sum of squares	Mean square	F
Personality	1	1609·88	1609·88	18·01
Diagnosis	5	109·93	21·99	—
Interaction	4*	223·84	55·96	—
Individuals	38	3396·50	89·38	—

* One cell is empty.

The personality × diagnosis analysis of variance presented in Table 3 leaves no doubt that the E scale is a measure of personality. There is no evidence, at least in this small sample, that it is related to type of illness or to any interaction between personality and diagnosis. Even if all the diagnostic variance were attributable to one comparison the resulting variance ratio would not be significant. This is not a proof that there is no such relation; it does, however, suggest that the relation cannot be very important, at least in the population of patients at Claybury Hospital. While it is reasonable to apply small-sample statistics to such comparisons as can be made on a random sample of patients within a single hospital it is not possible to estimate the parameters of a wider population unless an attempt has been made to sample mental hospitals.

The conclusion to be drawn from these findings is that the E scale is measuring personality and, after the extraction of this factor, is not related in any appreciable degree to diagnosis or the interaction of the two. Of course there is a possibility that it is related to other factors, for example sex, to some extent. Such further possibilities need to be explored using factorial techniques. The difficulty is that, with unequal numbers in the groups, the factorial analysis of variance with more than two factors becomes extremely complicated, but if numbers in the groups are equated there is a tendency to bias the sampling in an endeavour to make up the numbers in the rare categories.

A covariance analysis of the regression of E on HOQ confirmed the results of the analysis of variance. The regression lines within the diagnostic groups are not significantly different in slope ($F_{5,38} = 0{\cdot}55$) or in elevation ($F_{5,43} = 0{\cdot}40$).

The analysis of variance was performed by the method of fitted constants

(Snedecor, 1957, p. 388) which yields unbiased estimates of the means scores of the various groups. These are given in Table 4. No reliance can be placed on the estimates in the bottom half of the table. These six values are directly based on the scores of only thirteen patients. The remainder are probably a little more reliable. When interpreting this table it should be borne in mind that the basis of classification was a cutting point on another questionnaire.

Table 4. *Unbiased estimates of the mean scores of various personality and diagnostic groups on the MPI E scale*

(It is assumed that there is no interaction between personality and diagnosis.)

	Personality		
Diagnosis	Obsessoid	Hysteroid	Mean
Obsessional state	16·5	28·5	22·5
Anxiety state	19·5	31·5	25·5
Neurotic depression	17·9	29·9	23·9
Hysteria	19·4	31·4	25·4
Psychopathy	15·0	27·0	21·0
Personality disorder	15·2	27·3	21·2
Mean	17·2	29·3	23·3

Table 5. *Frequencies of patients in the various diagnostic and personality (obsessoid and hysteroid) categories*

		Sample	
Diagnosis	Personality	A	B
Obsessional state	O	1	1
	H	1	1
Anxiety state	O	16	19
	H	10	11
Neurotic depression	O	6	4
	H	3	2
Hysteria	O	3	2
	H	4	2
Psychopathy	O	2	2
	H	0	1
Personality disorder	O	1	3
	H	3	2
		50	50

PERSONALITY AND PATHOLOGY

If E score is unrelated to diagnosis why do hysteric groups almost invariably obtain higher scores than dysthymic groups? The answer lies partly, as Eysenck assumes, in the base rates of the two personality types in the subgroups of neurosis. Table 5 shows the frequencies of patients in the various categories in the present sample (labelled A) and in another sample obtained on exactly the same principles. Sample B is a group of consecutive admissions to the same hospital. The authors are not aware of any constraints which would account for the unusually close fit. Combining the two samples we find that the ratio of obsessoids to hysteroids is 1·7 : 1

in the dysthymic groups but only 1·1:1 in the hysteric groups. It is unlikely that these ratios are good estimates of the ratios in the neurotic population, since the hospital from which the samples are drawn has a strong bias to admit obsessoid rather than hysteroid patients (Martin & Caine, 1963). However, the direction of difference between the two ratios is not likely to be affected by this policy and it is sufficient to explain why hysterics obtain higher E scores than dysthymics. Hysterics are more extraverted than dysthymics in the sense that a hysteric is more likely than a dysthymic to have a hysteroid personality. However, this factor is at work in the present data and yet the difference between dysthymics and hysterics is not significant.

A further explanation is suggested by the work of Ingham & Robinson (1964). They took a group of patients diagnosed as hysterics and subdivided it into patients with conversion hysteria and patients with hysterical personality. The mean E score of the conversion hysterics was similar to that of dysthymics while the hysterical personalities obtained an extraverted score. The inference to be drawn is that psychiatrists are influenced by personality as well as by symptoms in assigning a diagnosis. This fact may account for the difference between the two ratios reported above. Considerations of sampling suggest that the confusion of personality and diagnosis may be present to a much greater extent in other validation studies. The sampling procedure which has been adopted in previous attempts to validate the MPI has not always been clearly stated. It may be surmised that only clear-cut cases were selected and that such sampling maximized the number of obsessoid dysthymics and hysteroid hysterics, and minimized the number of hysteroid dysthymics and obsessoid hysterics. If a psychiatrist expects a hysteric to have a hysteroid personality he will hesitate to diagnose an obsessoid person as a hysteric. Such a person is likely to be given a diagnosis such as 'mixed neurosis' and to be excluded from validation studies. No such residual category was employed in the present study. However, the relations between the frequencies shown in Table 1 may indicate that, in addition to its diagnostic employment, 'Anxiety State' was functioning as a residual category, and this may account for the unusually high mean score of patients with this diagnosis—a score which is close to the mean of the whole sample.

It may be asked why the authors prefer the unusual (and ugly) terms 'hysteroid' and 'obsessoid' to the popular 'extraverted' and 'introverted'. The former pair is historically prior to the latter and has an established place in the psychiatric literature. According to Mackinnon (1944) Janet was the first to use the hysteroid and obsessoid categories in the classification of the personalities of neurotics.

Jung (1917) first suggested an association between extraversion and the 'hysterical character'. He stated, however, that in the fully established illness of hysteria the patient becomes introverted. Introversion he associated with psychasthenia and dementia praecox. Again, however, at least in the case of dementia praecox, there is a reversal in the developed illness. The 'precocious dement' is an extravert.

Jung's conception of extraversion-introversion is one of direction of psychical energy, libido or interest. It has been widely supposed that certain personality traits are associated with direction of energy. Until this has been demonstrated it seems better to reserve 'extraversion-introversion' for Jung's dynamic dimension and to use 'hysteroid-obsessiod' for the dimension of personality traits measured by the HOQ and the MPI.

Brit. J. Psychol. (1964), 55, 4, *pp.* 453–462
Printed in Great Britain

AN OBSERVED AGE DIFFERENCE IN RESPONSES TO A PERSONALITY INVENTORY

By F. I. M. CRAIK

Medical Research Council Unit for Research on Occupational Aspects of Ageing, at the University of Liverpool

The two parts of the Heron Inventory give scores on two dimensions—emotional adjustment and sociability—which are held to be independent. However, results from a small sample study indicated that there was a strong positive association between the dimensions in a 'young' group but that no such relationship held in an 'old' group. This finding was confirmed for a group of 240 men ranging in age from 20 to 79. It was not confirmed in a similar group of 240 women.

In the male group, the result did not appear to be due to concomitant changes in non-verbal intelligence; it was apparently not caused by an increased tendency of the older subjects to agree or disagree with the items; it did not seem to be due to a change in the meaning or relevance of the items to older subjects, nor did it seem to be due to a decrease in the reliability of older subjects' responses. The result did, however, appear to be related to degree of adjustment. Possible reasons for this finding are discussed.

INTRODUCTION

The Heron Inventory (Heron, 1956) is a two-part personality measure giving scores on two dimensions—emotional adjustment and sociability—which are held to be independent. The Inventory consists of 110 statements, each of which must be checked 'True' or 'False' by the respondent. Of these statements, twenty are scored for adjustment and twelve for sociability (see Appendix). The remainder are designed to give the Inventory the flavour of a medical questionnaire.

The Inventory was given as part of a battery of personality and behavioural tests in a small-scale experiment on thirty-six male subjects of whom twelve were in their 20's, twelve in their 40's and twelve in their 60's. During analysis of the results it was observed that in the young group, emotional adjustment (H_1) correlated with sociability (H_2), $r = +0{\cdot}72$. This correlation was $+0{\cdot}61$ in the middle-aged group and $-0{\cdot}34$ in the old group. It was felt that the large difference in the relationship between H_1 and H_2 for the young and old groups might well be due to sampling error. On the other hand, if the result were reliable, it would affect the interpretation of the experimental results and might possibly have a bearing on personality changes with age. It was therefore decided to investigate the relationship between emotional adjustment and sociability, as measured by the Heron Inventory, at different age levels.

METHOD OF STUDY

The Inventory responses of 240 men were obtained from data available on the MRC Voluntary Research Panel (Heron, 1962). There were forty subjects in each of six 10-year age-groups, ranging in age from 20 to 79. The subjects, who were all volunteers, were drawn from a wide variety of occupations and each age-group covered a wide range of intelligence.

RESULTS

Product-moment correlations between H_1 (adjustment) and H_2 (sociability) were calculated within each decade group. The results are given in Table 1*a*. It is seen that a strongly positive relationship obtains in the 20's and that this decreases until there is a negligible relationship in the 60's and 70's. When the data are broken down further into 5-year groups of twenty subjects (Table 1*b*) the same pattern can still be seen.

Table 1. *Correlations between adjustment and sociability at different age levels*

(*a*) Male subjects						
Age range	20–29	30–39	40–49	50–59	60–69	70–79
$r_{H_1H_2}$	+0·62	+0·44	+0·31	+0·41	+0·01	−0·05
	(N = 40 for each group)					
(*b*) Male subjects						
Age-group	20	25	30	35	40	45
$r_{H_1H_2}$	+0·66	+0·63	+0·59	+0·25	+0·40	+0·19
Age-group	50	55	60	65	70+	70+
$r_{H_1H_2}$	+0·42	+0·40	+0·28	−0·21	0·00	−0·08
	(N = 20 for each group)					
(*c*) Female subjects						
Age range	20–29	30–39	40–49	50–59	60–69	70–79
$r_{H_1H_2}$	+0·16	+0·44	+0·14	+0·52	−0·01	+0·42
	(N = 40 for each group)					
(*d*) Male subjects						
Age range	20–29	30–39	40–49	50–59	60–69	70–79
Mean Raven scores (high)	54·8	52·2	53·6	48·6	44·5	37·7
$r_{H_1H_2}$	+0·66	+0·52	+0·18	+0·41	+0·26	−0·06
Mean Raven scores (low)	44·7	44·0	44·1	38·8	29·2	17·6
$r_{H_1H_2}$	+0·59	+0·35	+0·44	+0·41	+0·03	+0·07
	(N = 20 for each group)					

The significance of the effect which age has on the relationship between H_1 and H_2 can be assessed roughly by calculating a Spearman rank-difference correlation between age-group and $r_{H_1H_2}$. This was done for the data in Table 1*b* and a value of ρ equal to $-0{\cdot}82$ ($P < 0{\cdot}01$) was obtained. Thus the tendency for $r_{H_1H_2}$ to decrease with age is statistically significant.

A more powerful method of assessing the significance of this trend is to examine the regressions of H_1 on H_2 and H_2 on H_1. If age has the predicted effect on $r_{H_1H_2}$ then the slopes of the regression lines should decline with age from positive to zero, while at the same time, the intercepts of the regression lines should increase with age. The significance of these effects was examined by multiple correlation treating age as a continuous variable.

The relationship between H_1 and H_2 can be specified by an equation of the form $H_1 = a + bH_2$. A similar relationship holds between age and a, and between age and b. Thus (where A = age) we can substitute $(a_0 + b_0A)$ and $(a_1 + b_1A)$ for a and b respectively. That is,

$$H_1 = a + bH_2.$$

Also $$a = (a_0 + b_0A) \quad \text{and} \quad b = (a_1 + b_1A),$$

therefore $$H_1 = (a_0 + b_0A) + (a_1 + b_1A)H_2.$$

Values for a_0, b_0, a_1 and b_1 can be obtained by fitting a multiple regression plane to the data. The obtained equations were:

$$H_1 = (3{\cdot}86^* + 0{\cdot}018A) + (0{\cdot}999^{**} - 0{\cdot}0123A^*)H_2,$$
$$H_2 = (-0{\cdot}43 + 0{\cdot}080A^{**}) + (0{\cdot}709^{**} - 0{\cdot}0083A^*)H_1,$$

where A = age; $* = P < 0{\cdot}05$; $** = P < 0{\cdot}01$.

The regression coefficients have the expected signs (i.e., as age increases, a increases but b decreases in value) and are significantly removed from zero.[†]

Thus again it can be concluded that there is a statistically significant tendency for the strength of the relationship between H_1 and H_2 to decrease with age. That is, the slight positive correlation between the dimensions which appears if a large sample of mixed ages is taken conceals a relationship which changes progressively from highly positive to zero.

It should be pointed out here that the foregoing analysis applies only to male subjects. The results for 240 female subjects are given in Table 1*c*. It is seen that while there is again a slight positive correlation between H_1 and H_2 at most age levels there is apparently no trend with age.

FURTHER ANALYSIS OF DATA

Before attempting to discuss the meaning and possible implications of the finding on the 240 male subjects it was thought advisable to rule out several factors which might have given rise to it.

A recent study (Edwards & Wine, 1963) reported an apparent decline with age of fourteen personality and biographical variables, but when non-verbal intelligence scores (on Raven's Progressive Matrices) were taken into account the decline with age became insignificant in all but four cases. As non-verbal intelligence, measured by Progressive Matrices, declines with increasing age in this present sample, it is possible that non-verbal intelligence, and not age as such, was affecting the relationship between emotional stability and sociability.

The mean Matrices scores were calculated for each of the twelve 5-year age-groups cited in Table 1*b*. Three correlation coefficients were calculated using Kendall's τ; age with $r_{H_1H_2}$ $(-0{\cdot}67)$; age with mean Matrices scores $(-0{\cdot}76)$ and Matrices scores with $r_{H_1H_2}$ $(+0{\cdot}48)$. The partial τ between age and $r_{H_1H_2}$, holding Matrices scores constant was also calculated. This gave a value of τ equal to $-0{\cdot}54$. It is not yet possible to test the significance of this value but it is seen that when the effect of non-verbal intelligence is removed a substantial, though reduced, correlation between age and $r_{H_1H_2}$ remains.

Further, the forty subjects in each 10-year group were divided into the twenty highest scorers and the twenty lowest scorers on Raven's Matrices. Product-moment correlations between emotional stability and sociability were calculated separately for the 'High Matrices' group and the 'Low Matrices' group at each age level (Table 1*d*). In three of the six cases the 'High Matrices' group has the larger correlation, one pair has the same correlation and in two cases the 'High Matrices' group has the lower correlation. In no case is the difference between the pairs of correlation coefficients significant. This further suggests that non-verbal intelligence is not the critical factor.

[†] This method was suggested and the calculations carried out by Mr Frank Brechling of the National Economic Development Office.

The responses made to the scored items of the Inventory were then examined. As the major differences in the $H_1.H_2$ correlations were between the 20's and 30's on one hand and the 60's and 70's on the other, the responses of these two groups only were considered. Thus there were eighty subjects in each of two groups: 'Young' and 'Old'.

The possibility that older subjects might have an increased tendency to agree or disagree with items regardless of their content was first examined, but no such increased tendency was found. (Of the thirty-two scored items, the mean number answered 'True' was 14·9 for the Young group and 14·1 for the Old group. The difference between the means was found not significant at the $P = 0{\cdot}05$ level by the t test.)

It was also considered possible that certain statements in the Inventory might have a different meaning or relevance to older subjects. For example, the response 'False' to the statement: 'At a gay party I can usually let myself go and have a thoroughly good time' might be considered a good indicant of unsociability in a younger person but might well be irrelevant to that quality in an older person. The numbers of subjects in each group (Young and Old) who answered each item in the scored direction were examined for significant differences by χ^2. Of the twenty items scored for emotional adjustment, eight showed significant differences between the Young and Old groups. Of these eight, five items (7, 20, 44, 46 and 61) were answered significantly more often by the Young group and three items (50, 55 and 69) were answered significantly more often by the Old group. Of the twelve items scored for sociability, three showed significant differences between the two groups. One (82) was answered significantly more often by the Young group and two items (97 and 101) were answered significantly more often by the Old group.

Thus, while there *are* differences between the Young and Old groups in the items going to make up H_1 and H_2 scores, it is difficult to see that the 'emotional maladjustment' items answered more often by the Young group are particularly 'unsociable' in character. This feeling is reinforced by the fact that when all items answered differently by the Young and Old groups were removed and amended scores for H_1 (out of 12) and H_2 (out of 9) were calculated for each subject, the correlations between H_1 and H_2 still showed an age difference:

Age-group	20	30	60	70
$r_{H_1H_2}$ (amended scores)	+0·54	+0·40	−0·07	−0·13

It thus seems unlikely that the observed difference between the Young and Old groups in the $H_1.H_2$ correlation is an artefact due to a difference in meaning of the questionnaire items for younger as compared with older subjects.

The reliabilities of the two parts of the Inventory were calculated for the Young and Old groups separately by the Kuder–Richardson method. For the emotional adjustment scale these reliabilities were 0·67 and 0·72 for the Young and Old groups respectively. The reliability coefficients for the sociability scale were 0·72 and 0·62 respectively. These values are slightly lower than those obtained by Heron (1956) but are reasonable in view of the smaller numbers involved. The relative consistency of the reliability coefficients between Young and Old groups is held to discount the possibility that old subjects show less reliability than younger subjects in completing the questionnaire. The decrease with age in the correlations between H_1 and H_2 cannot therefore be due to this factor.

A further possibility seemed to be that there was a systematic change with age in the proportions of adjusted to maladjusted subjects. Heron (1956) suggests that subjects can be classified by their scores on H_1 as: 'probably well-adjusted' (0–7); 'doubtful' (8–9); or 'probably maladjusted' (10 +). As approximately half of the subjects gave scores of 8 and over it was decided to take this score as a cutting point. In an attempt to gain clarity of exposition, subjects scoring 0–7 will be referred to as 'well-adjusted' and those scoring 8 + as 'poorly-adjusted', although it should be stressed that this classification is a relative one.

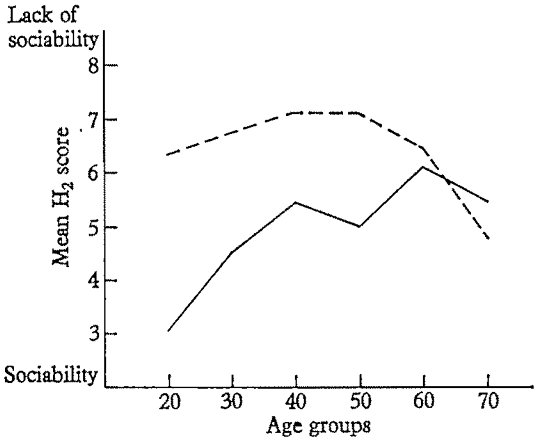


Fig. 1. Mean H_2 scores plotted separately for 'well-adjusted' and 'poorly-adjusted' subjects at each age level. – – – poorly-adjusted subjects; —— well-adjusted subjects.

The number of subjects in each category is given in Table 2. The 70's group has a high number of well-adjusted subjects but the distribution of the 60's group is practically identical with that of the 20's and 30's. The tendency for the number of well-adjusted subjects to increase with age was found non-significant using Kendall's τ. However, when mean sociability scores are plotted separately for the well- adjusted and poorly-adjusted subgroups, it is seen (Fig. 1) that in the 20's well-adjusted subjects are markedly more sociable. The discrepancy between the mean H_2 scores decreases with age, however, until in the 70's the poorly-adjusted subgroup is slightly more sociable.

The $H_1 . H_2$ correlations *within* each subgroup were then examined. The results are given in Table 2. It is seen that in the well-adjusted subgroups a low positive correlation generally obtains between adjustment and sociability and there appears to be no regular tendency for this to change with age. Within the poorly-adjusted subgroups, there is a similar slight positive correlation in the 20's, 30's and 40's, but there is a slight negative correlation between the dimensions in the 50's, 60's and 70's. That is, in younger subjects greatest maladjustment is related to unsociability while for older subjects there is a tendency for greatest maladjustment to be related to sociability.

Table 2. *Rank-difference correlations between adjustment and sociability within subgroups*

Age-group	Well-adjusted subjects	Poorly-adjusted subjects
20	+0·26 ($N = 21$)	+0·37 ($N = 19$)
30	+0·12 ($N = 21$)	+0·28 ($N = 19$)
40	−0·17 ($N = 25$)	+0·31 ($N = 15$)
50	+0·40 ($N = 24$)	−0·21 ($N = 16$)
60	+0·26 ($N = 22$)	−0·42 ($N = 18$)
70	+0·03 ($N = 28$)	−0·14 ($N = 12$)

(Rank-difference correlation coefficients were calculated as N's were small and distributions not normal.)

Thus in the 20's and 30's there is a positive correlation between stability and sociability within both subgroups while the high adjustment groups are high on sociability and the low adjustment groups are relatively low on sociability. For the whole Young group a positive correlation between stability and sociability would thus be expected. In the 60's and 70's, however, there is a negative relationship between stability and sociability in the poorly-adjusted groups while the high and low adjustment groups differ little in their mean sociability scores. Thus a negligible correlation between stability and sociability in the total Old sample would be expected and is obtained.

These differences between the well-adjusted and poorly-adjusted subgroups go some way towards 'explaining' the overall finding or at least make it necessary for further examination to be at a slightly deeper level of analysis.

DISCUSSION

In his description of the Heron Inventory, Heron (1956) states: 'The term "sociability" has been selected for this [second] part of the present inventory mainly because it is somewhat less likely than is "introversion-extraversion" to be narrowly identified with any particular school of thought in contemporary psychology.' Thus Heron appears to view 'sociability' as being fairly closely related to 'introversion-extraversion' and this is taken as justification for citing studies which discuss the latter concept.

For the 240 male subjects the age trend described above appears to depend on three factors: the different age trends in sociability scores for the well-adjusted and poorly-adjusted groups; the differences between the well-adjusted and poorly-adjusted groups in $H_1.H_2$ correlations within subgroups and thirdly, the large difference in sociability scores between the well-adjusted and poorly-adjusted groups for the Young subjects and the lack of such a difference for Old subjects. These three factors will be discussed in turn.

Several studies (e.g. Bendig, 1960; Calden & Hokanson, 1959) have shown that introversion, as measured by questionnaire, increases with age. The age trend on sociability for well-adjusted subjects (Fig. 1) is thus the expected result, except for falls in the 50's and 70's. The drop in the 50's is probably a sampling error as a further group of seventy well-adjusted men in their 50's gave a mean H_2 score of 5·6. The slight drop in the 70's may again be a sampling error, but the fact that for poorly-adjusted subjects both the 60's and the 70's show such a drop suggests that the observations may be valid. If this were the case, it might mean that after a certain age (slightly younger in a poorly-adjusted group) subjects prefer to answer the questionnaire items in the direction of greater sociability. This preference is understandable if sociability is considered a socially desirable quality. The possibility that social desirability of response does become a more important factor with age is strengthened by the fact that the mean scores on emotional maladjustment are lowest in the 60's and 70's. Alternatively, it is possible that in the older age-groups, only the more sociable and stable members of the population would volunteer for psychological testing. If this factor is operating, then Fig. 1 shows that the tendency towards lack of sociability with increasing age is biased in the 60's or 70's by the tendency for volunteer samples to contain an increasingly greater number of sociable or extravert subjects. Further, that this bias operates more strongly and at an earlier age in a poorly-adjusted group.

Carrigan (1960) has reviewed the literature on extraversion-introversion as a dimension of personality. One of the major issues which she considers is the independence of extraversion-introversion and adjustment. She concludes that a clear-cut answer cannot be given but also states: 'In many instances, however, the characteristics associated with "introversion" continue to have a strong maladjusted flavor.' Other studies (Bernreuter, 1934; Vernon, 1938; Eysenck, 1959) have found a low but persistent relationship between extraversion and emotional adjustment. A recent paper by Eysenck & Eysenck (1963) suggests that the factor of extraversion is made up of two main traits; sociability and impulsiveness. Further, that adjustment is correlated slightly negatively with impulsiveness but slightly positively with sociability. From this evidence, a slight positive correlation between sociability and emotional stability would be expected and this is in fact found for the well-adjusted subgroups at all ages (except the 40's) and for the poorly-adjusted subgroups in the 20's, 30's and 40's. The correlations between H_1 and H_2 for the poorly-adjusted subgroups in the 50's, 60's and 70's are all negative, and although none reaches significance some possibility remains that for these groups sociability is related to maladjustment.

In discussing the third factor—the large difference in sociability scores between well-adjusted and poorly-adjusted groups for young subjects and the lack of such a difference for old subjects—the overall correlation between H_1 and H_2 must be considered. Broadbent (1958), discussing the independence of neuroticism and introversion, points out the possibility of cultural differences in this connexion. He states: 'In some societies introversion may well be more synonymous with social abnormality and so with neuroticism than it is in others.' Also, Carrigan (1960) states '...in a culture such as ours which places a high premium on interpersonal interaction, the characteristic avoidance of such interaction—associated with social introversion [i.e. lack of sociability]—might be considered maladaptive.'

It seems possible that lack of sociability might well be associated more strongly with social abnormality and emotional maladjustment in young subjects. Thus age differences might take the place of Broadbent's suggested cultural differences and the decreasing magnitude of the $H_1.H_2$ correlations reflect the declining importance with age of social interaction. Put another way it is possible that young people see sociability as being an integral part of adjustment while for older persons the two factors are seen as much less closely related.

The main objection to accepting the findings of this study as an important age difference on the Heron Inventory is the fact that the trend is absent in female subjects. It is certainly possible that the female result is the true one and the male result a 'Type I' error although the levels of significance reported in the Results section make this appear unlikely. When the female scores are subjected to the same type of analysis as those of the males into well-adjusted and poorly-adjusted groups the results, briefly, show that much the same age trends on sociability exist for the well-adjusted and poorly-adjusted groups as for the male groups, secondly that the $H_1.H_2$ correlations within subgroups are all generally slightly positive and thirdly that while the poorly-adjusted group is higher than the well-adjusted group on H_2 at all age levels except the 60's there is no marked difference between the groups as in the male 20's and 30's. These differences between the male and female samples may again be attributable to sampling error or there may be a genuine sex difference. If the latter is the case the most likely explanation seems to be that social interaction is either less important for young females or at least sociability is not seen by them as an integral component of adjustment.

Turning again finally to the male sample, it should be stressed that the results discussed above show *differences* between age-groups *rather than changes with age.* This proviso applies, of course, to any cross-sectional age study. It is important to point out, however, that while it is certainly possible, as suggested above, that attitudes change with increasing age, it is also possible that the accepted attitudes held by society are changing and that the importance of social interaction is greater today than it was 40 years ago. In this case it might be that the subject's cultural experience is more important than his age in this connexion.

If the trend reported here is the true result the finding is important for the study of personality changes with age. It may mean that the structure of personality alters with age or at least that older persons project their personalities differently when measured on tests very often standardized on a younger group. The subsidiary finding that similar scores on such variables as adjustment and sociability may be compounded of different items, though its relevance was discounted for the present discussion, may prove important in the assessment of what the scores actually 'mean'. Kuhlen (1959) remarks with regard to ageing: 'Capacities change, problems and habits change and the mechanism utilised for adjustive purposes likely vary in incidence with age.' Thus the questions best suited to elicit scores reflecting adjustment or sociability may also change with the age of the respondents or if the same questions are used the significance of the responses may change.

The main issue which emerges from this discussion is whether the observed result in fact reflects progressive changes in personality structure with age or whether it was due to sampling error, the Inventory used, or differences in cultural experience. Further work with different samples, differently obtained samples and different methods of measuring emotional stability and sociability must be undertaken to ascertain which of these possibilities is giving rise to the observed effect.

The author wishes to thank Dr Alastair Heron and his colleagues in the Medical Research Council Unit for Research on Occupational Aspects of Ageing, and Prof. L. S. Hearnshaw and Dr D. B. Bromley of the Department of Psychology, University of Liverpool, for many helpful comments and suggestions during the preparation of this paper.

APPENDIX

List of items used as basis for scoring in the Heron Inventory

		S.R.
First Part (emotional maladjustment)		
3.	I sweat very easily even on cool days	T
5.	I am troubled by sick headaches	T
6.	My feelings are easily hurt	T
7.	Sometimes when I am not feeling well I am cross	T
9.	I consider myself rather a nervous person	T
15.	I never have nightmares	F
20.	I wake up fresh and rested most mornings	F
23.	My eyesight is sometimes blurred	T
25.	I worry over possible misfortunes	T
27.	It is difficult to irritate me	F
33.	My bowels are not opened every day	T
44.	My physique is excellent	F
46.	Once in a while I think of things too bad to talk about	T
47.	I often feel self-conscious in the presence of superiors	T
50.	I feel that very little works out the way it should in life	T
55.	I have no particular physical or health problems	F
60.	I have periods of such great restlessness that I cannot sit long in a chair	T
61.	My hands often feel damp and clammy	T
69.	I catch cold frequently	T
72.	I often jump to conclusions	T
Second Part (sociability)		
79.	The presence of the opposite sex makes me feel shy	T
82.	What others think of my actions does not bother me much	F
83.	I sometimes treat my work as if it were a matter of life and death	T
84.	I like to play practical jokes on people	F
86.	I find it easy to act naturally at a party	F
93.	I like to have time to be alone with my thoughts	T
97.	I like to have a lot of social engagements	F
101.	I tend to be over-conscientious	T
102.	I am usually a good mixer	F
103.	I think of myself as a lively sort of person	F
104.	When I am bored I feel like stirring up some excitement	F
109.	At a gay party I can usually let myself go and have a thoroughly good time	F

S.R. Significant response.

Note. A high score on H_1 indicates maladjustment; a high score on H_2 indicates *lack* of sociability.

REFERENCES

BENDIG, A. W. (1960). Age differences in the interscale factor structure of the Guilford Zimmerman Temperament Survey. *J. Consult. Psychol.* **24**, 134–8.

BERNREUTER, R. G. (1934). The imbrication of tests of introversion-extroversion and neurotic tendency. *J. Soc. Psychol.* **5**, 184–201.

BROADBENT, D. E. (1958). *Perception and Communication.* London: Pergamon Press.

CALDEN, C. & HOKANSON, J. E. (1959). The influence of age on MMPI responses. *J. Clin. Psychol.* **15**, 194–5.

CARRIGAN, P. M. (1960). Extraversion-introversion as a dimension of personality: a reappraisal. *Psychol. Bull.* **57**, 329–60.

EDWARDS, A. E. & WINE, D. B. (1963). Personality changes with age: their dependency on concomitant intellectual decline. *J. Gerontol.* **18**, 182–4.

Eysenck, H. J. (1959). *Manual of the Maudsley Personality Inventory.* London: University of London Press.

Eysenck, S. B. G. & Eysenck, H. J. (1963). On the dual nature of extraversion. *Brit. J. Soc Clin. Psychol.* **2**, 46–55.

Heron, A. (1956). A two-part personality measure for use as a research criterion. *Brit. J. Psychol.* **47**, 243–51.

Heron, A. (1962). The Liverpool functional age project: preliminary communication on population and methodology. *Experientia*, **18**, 470.

Kuhlen, R. G. (1959). Aging and life adjustment. In *Handbook of Aging and the Individual* (ed. by J. E. Birren). Chicago: University of Chicago Press.

Vernon, P. E. (1938). *The Assessment of Psychological Qualities by Verbal Methods.* M.R.C. Industrial Health Research Board, Report no. 83. London: H.M. Stationery Office.

(*Manuscript received* 31 *October* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 463–468
Printed in Great Britain

RESPONSES BY THE FERRET TO STIMULUS CHANGE

BY R. N. HUGHES
University of Canterbury, New Zealand

Ten ferrets were allowed to move about freely in a Y-maze with an additional centre arm, for three periods of time. A change in brightness was then effected by removal of an anomalous arm insert and entrance archway to produce a maze of uniform brightness. On a subsequent critical trial the ferrets were observed to enter first the arm in which this change had been made. The results are discussed in terms of perceptual satiation, development of a response habit, response inhibition and stimulus change. The last possibility is shown to be the most appropriate, thus confirming conclusions from rat studies.

I. INTRODUCTION

It has been suggested by Dember (1956) that the results of the experiment by Kivy, Earl & Walker (1956) on spontaneous alternation in white rats could be explained by the 'response to novelty' hypothesis rather than by 'stimulus-satiation'. Kivy *et al.* exposed rats for three different periods of time (1, 15 and 30 min.) to a T-maze in which the subject was prevented from entering the goal arms, but was still able to see into them through a glass partition across each arm entrance. The arms were both either entirely black or entirely white. Immediately after an exposure a second trial was carried out for which the partitions were removed and one of the arms changed to the alternate brightness. For the 15 and 30 min. exposures the rats entered the changed goal arm a significantly greater number of times. These results have been interpreted as evidence for 'stimulus-satiation'. That is, the rat on its second trial avoided the stimulus-brightness for which it was perceptually satiated on its first trial.

Dember repeated the experiment of Kivy *et al.* with one major alteration. On each exposure (which was of 15 min. duration), one arm was white and the other black, thus presumably satiating the rats for both black and white. But on each second trial one arm was changed to the same brightness as the other. Dember observed a significant number of entries of the arm that had changed from what it was during exposure, and thus concluded that the rats had responded to the 'novel' stimulus, an explanation which would also account for the results of Kivy *et al.*

The present study was designed to test Dember's findings by observing a different species in a slightly more complicated apparatus, and allowing free choice on both exposure and re-exposure trials. Basically, ferrets were allowed to move about freely for a fixed period of time in a three-arm Y-maze with one arm of opposite brightness to the rest of the apparatus (i.e. an exposure trial). First arm entries were observed after the different arm had been changed to the same brightness as the rest of the apparatus (i.e. a critical trial). Such freedom of movement on exposure trials ensured that the ferret was not restrained from locomotor activity within the arms, being able to visit each one according to its own relative preferences. The experiment consisted of two parts. For each exposure trial, Part I was carried out in a white maze interior with one black arm, and Part II in a black maze interior with one white arm.

II. METHOD

Subjects

The subjects were ten trapped, wild New Zealand ferrets (*Mustela putorius*). Subjects 1–5 were used in Part I of the experiment, and subjects 6–10 were used in Part II. Frequent handling for some 3 months prior to the experiment had tamed all subjects sufficiently for testing.

Ferrets were considered as subjects mainly because their level of 'exploratory drive' (Berlyne, 1950) is extremely high, which makes them ideal subjects for any experiments in which exploration is not discouraged. As it was desirable to see whether conclusions on alternation behaviour drawn from results with rats would hold for another species, it was felt that ferrets were suitable animals to use because of their 'curiosity', their smallness in comparison with other carnivores, and their general shape, which is somewhat similar to that of rats.

Apparatus

The apparatus consisted of a conventional enclosed-alley Y-maze unit with an additional arm (see Fig. 1). The two outer arms were each set at an angle of 45 degrees from the third arm. Each arm was framed by a removable hardboard archway of the same height and width as the arm, but with a rectangular hole 10 in. high and 6 in. wide. Its function was merely to emphasize the brightness of the arm behind it. There was one spare archway of the same brightness as the rest of the maze interior for the critical trials. Changed brightness of an arm was effected by means of an archway and plywood insert of opposite brightness to the rest of the apparatus. This insert consisted of rear and side walls and a floor, and occupied the full internal area of any arm.

The starting box was covered with a hinged lid and its distal end was bounded by a hardboard guillotine slide. Its interior was painted mid-grey.

When an exposure or critical trial was in progress, the choice area and arms were covered by a wire-netting frame. The stem was covered by a piece of heavy duty wire-netting. These covers were capable of being quickly removed to enable removal of the insert and replacement of the accompanying archway after each exposure trial.

Normal room illumination was maintained throughout by means of two 5 ft. fluorescent tubes suspended approximately 6½ ft. above the apparatus.

To facilitate the observations and operations, and also to reduce extra-maze visual cues, the apparatus was placed on the floor of a 6 ft. 10 in. square open-field maze, which was raised 16 in. off the floor of the room, and had the stem of the apparatus protruding from one corner where it was supported by a metal stand.

Procedure

Part II (black maze interior) was not started until Part I (white maze interior) had been completed. Three periods of time were used for the exposure trials, i.e. 1, 5 and 10 min. These times and the position of the different arm were pre-determined in an irregular fashion for each trial, so that each subject received four trials at each exposure time, with the altered arm in two positions once and in one position twice. The subjects in both parts received one exposure and one critical trial per day for a total of 12 days.

In an effort to control any possible effects of the hunger variable, all subjects were allowed free access to their basic diet, horseflesh and milk, 1 hr. before each experimental session. The first subject in Part I was then placed in the starting box. After approximately 30 sec. the guillotine slide was raised, thus allowing the animal to move into the maze proper. It was kept there and allowed to move about freely in the apparatus for the duration of the appropriate exposure time and then removed and returned to the starting box. The black insert was then removed and the related black archway replaced by one of the same brightness as the rest of the apparatus—i.e. white. This operation never took more than 35 sec. nor less than 20 sec. Then the starting box slide was raised again and the subject's first body-length entry of an arm through the archway recorded. The latency between raising of the slide and first entry was approximately 20 sec. At no time did a subject fail to enter an arm. Following an entry the subject was removed from the apparatus and in an effort to reduce olfactory cues for the next subject, the whole interior was mopped out with a strong smelling disinfectant solution, and liberally sprayed with a household deodorizer. Exactly the same procedure was used for the other subjects in Part I.

The procedure was the same for both parts, with the maze white and the insert and related archway black on exposure trials (and white on critical trials) for Part I, and with the maze black and the insert and related archway white on exposure trials (and black on critical trials) for Part II.

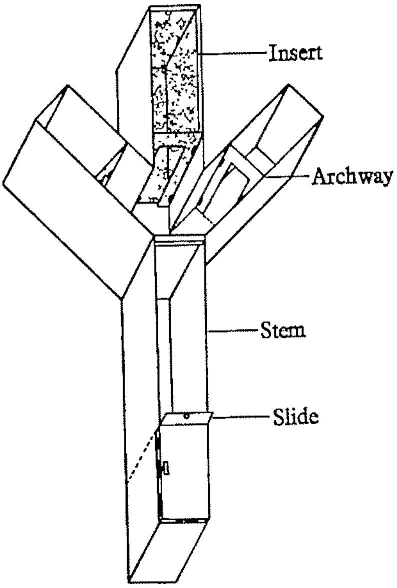


Fig. 1. Apparatus with covers removed (drawn to scale approx. 1:22), as for an exposure trial of Part I.

III. RESULTS

Total first entries of the altered arm on critical trials ('responses to change') for each subject after each exposure time are recorded in Table 1.

As the basic observations were proportions these total 'responses to change' were transformed by the arcsine transformation.* This was necessary to stabilize the variances (Winer, 1962, p. 221). An analysis of variance for repeated measures (Kogan, 1948) was then performed on the transformed scores.

'Maze conditions' ($F = 2{\cdot}65$), 'exposure times' ($F = 2{\cdot}22$) and 'conditions × ex-

Table 1. *Responses to change in four critical trials by each ferret after each exposure time*

	Part I					Part II				
Subject ...	1	2	3	4	5	6	7	8	9	10
1 min.	0	2	0	3	1	4	2	4	3	2
5 min.	3	3	2	3	1	4	2	4	4	2
10 min.	2	4	2	2	3	4	3	3	3	1

---

* These data were all transformed $Y = \sin^{-1}\surd(D/N)$; where D (responses to change) $= 0$, $Y = \sin^{-1}\surd(1/2N)$; where $D = N$ (trials), $Y = \sin^{-1}\surd[1-(1/2N)]$ (L. E. Moses, 1963, personal communication to a colleague of the writer).

posure times' ($F = 2{\cdot}39$) were not significant, but 'subjects' ($F = 3{\cdot}47$) was significant at the 0·025 level.

Using M.S. for 'subjects' (364·07) as an estimate of S^2, and 8 degrees of freedom, t values for the difference between observed transformed means and that predicted by chance (i.e. 35·2°), were calculated for 'responses to change' after each exposure time in each part. No result in Part I was significant (t, 1 min. $= 0$; t, 5 min. $= 1{\cdot}83$; t, 10 min. $= 2{\cdot}05$). However, all three results in Part II were significant at the 0·05 level (t, 1 min. $= 2{\cdot}61$; t, 5 min. $= 2{\cdot}84$; t, 10 min. $= 2{\cdot}40$).

IV. DISCUSSION

The results of critical trial responses clearly indicate that the ferrets had entered first an arm that had changed in brightness from what it was during exposure. It can thus be assumed that they were induced to approach the changed brightness rather than to avoid the unchanged brightness, an assumption not in accordance with satiation theory. Obviously each subject would have been satiated for both brightnesses on the exposure trials (especially the unchanged brightness) so that it should show no significant preference for entering first any particular arm on the critical trials.

Table 2. *Total entries of two black arms and one white during exposure trials by each ferret of Part II*

	1 min.		5 min.		10 min.	
Subject	Black	White	Black	White	Black	White
6	11	5	38	22	38	21
7	10	6	32	18	47	24
8	10	4	33	16	58	26
9	16	10	42	18	84	46
10	9	4	31	15	44	23
Total	56	29	176	89	271	140

It might be argued that the results were due to satiation for a 'stimulus-in-a-place'. Apart from such an explanation being less parsimonious than 'response to change' it has been shown by Walker, Dember, Earl & Karoly (1955) that the stimulus and its position are independent factors in alternation situations, and that 'stimulus' supersedes 'place' in relative importance.

Another possibility is that the ferrets had merely acquired a habit of responding to an arm that was preferred in the exposure trials because of its 'oddity'. With the number of entries of each arm used as a measure of relative preference for them, it is obvious from Table 2 that the ferrets of Part II had shown no preference for the 'odd' arm. (Unfortunately these data were not recorded for Part I subjects.)

It is therefore unlikely that critical trial choices would have been influenced by the 'oddity' situation of the exposure trials.

The only other likely explanation is in terms of inhibition for a previous turning response on a second trial (Solomon, 1948). But in view of the fact that subjects were not restricted to one arm-entry on exposure trials, so that proprioceptive stimulation was more or less randomly received, this would be difficult to accept. Further, Montgomery (1952) and Glanzer (1953) have ably demonstrated that stimuli and not responses are the relevant alternated factors in alternation experiments. Their findings have subsequently been confirmed by other workers (e.g. Walker *et al.* 1955).

Contrary to the findings of Kivy *et al.* (1956) time of exposure apparently had no effect on critical trial choices. Whether this is due to the different species or the different procedure can only be left to speculation at this stage. The apparatus may have been too small, thus causing rapid adaptation regardless of length of exposure, or the exposure times may have been subjectively too alike to cause any differential effects on preferences of the subjects.

Although not statistically significant, there was a marked suggestion of a poorer performance in the white maze of Part I. If on further experimentation with more subjects a difference between white maze and black maze responses could be demonstrated, it would probably be best accounted for by brightness aversion. The ferret is a nocturnal animal showing aversion to high illumination, which tends to disrupt its visual orientation in a given environment (McCann, 1955). Consequently, total whiteness of a maze, further intensified by fluorescent lighting could be expected to increase the difficulty of visually detecting a changed arm. Similar white aversion tendencies in exploration by rats have been suggested by Williams & Kuchta (1957).

The general conclusion from this experiment is that even when they had three homogeneous alternatives to choose from the ferrets had responded to the changed stimulus by approaching and entering first the maze arm which had been altered in brightness from what it was during prior exposure. It is interesting to note that Levine, Staats & Frommer (1958) failed to demonstrate response to change in rats using a Y-maze and procedure similar to those of Dember & Millbrook (1956). However, Dember (1958) later pointed out that he and Millbrook had used 3 min. exposure (as opposed to 15 min. by Levine *et al.*) because the stem of the Y-maze was shorter than the stem of the T-maze used in the original experiment (Dember, 1956), meaning that the rats would be more confined thus increasing emotionality with long exposures. He accordingly repeated the experiment of Levine *et al.* with 3 min. exposure, and observed significant responses to change. Levine (1958) also repeated the experiment with 3 min. exposure but failed to get significant results. However, Walk (1960) managed to replicate Dember's results and suggested that Levine's 'failure of replication may be due to a difference in the extra-maze environment of the two laboratories'. The results of the present experiment would also tend to support Dember's conclusions, thereby increasing the possibility of stimulus change being a directing variable in the locomotor activity of many organisms by showing its function in a *feral* species far removed, phylogenetically, from the white rat.

As avoidance of a familiar arm brightness has been shown to be unacceptable as an explanation, it is further concluded that it was specifically 'novelty' generated by the attention-arousing change in brightness that had initiated *approach* by the ferrets of the altered arm.

REFERENCES

BERLYNE, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.* **41**, 68–80.

DEMBER, W. N. (1956). Response by the rat to environmental change. *J. Comp. Physiol. Psychol.* **49**, 93–5.

DEMBER, W. N. (1958). A comment on 'studies on "response by the rat to environmental change"'. *Psychol. Rep.* **4**, 242.

DEMBER, W. N. & MILLBROOK, B. A. (1956). Free choice by the rat of the greater of two brightness changes. *Psychol. Rep.* **2**, 465–7.

GLANZER, M. (1953). The role of stimulus satiation in spontaneous alternation. *J. Exp. Psychol.* **45**, 387–93.

KIVY, P., EARL, R. W. & WALKER, E. L. (1956). Stimulus context and satiation. *J. Comp. Physiol. Psychol.* **49**, 90–2.

KOGAN, L. S. (1948). Analysis of variance-repeated measurements. *Psychol. Bull.* **45**, 131–43.

LEVINE, S. (1958). A reply to a comment by Dember. *Psychol. Rep.* **4**, 433.

LEVINE, S., STAATS, S. R. & FROMMER, G. (1958). Studies on 'response by the rat to environmental change'. *Psychol. Rep.* **4**, 139–44.

MCCANN, C. (1955). Observations on the polecat (*Putorius putorius* Linn.) in New Zealand. *Rec. Dom. Museum*, **2**, 151–65.

MONTGOMERY, K. C. (1952). A test of two explanations of spontaneous alternation. *J. Comp. Physiol. Psychol.* **45**, 287–93.

SOLOMON, R. L. (1948). The influence of work on behaviour. *Psychol. Bull.* **45**, 1–40.

WALK, R. D. (1960). Response of dark- and light-reared rats to stimulus change. *J. Comp. Physiol. Psychol.* **53**, 609–11.

WALKER, E. L., DEMBER, W. N., EARL, R. W. & KAROLY, A. J. (1955). Choice alternation. I. Stimulus *vs.* place *vs.* response. *J. Comp. Physiol. Psychol.* **48**, 19–23.

WILLIAMS, C. D. & KUCHTA, J. C. (1957). Exploratory behaviour in two mazes with dissimilar alternatives. *J. Comp. Physiol. Psychol.* **50**, 509–13.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

(*Manuscript received* 12 *November* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 469–476
Printed in Great Britain

A DEUTERANOMALOUS ARTIST*

By R. W. PICKFORD
Psychology Department, University of Glasgow

An artist, who claimed to be colour blind, according to a report in the local newspapers, kindly co-operated in efficient colour vision tests, two long discussions about his paintings with the writer, and allowed nine of his paintings to be photographed as colour transparencies. This paper presents notes made during these discussions, detailed results of the tests, and a short commentary on his paintings.

THE ARTIST

The artist, Mr Donald R. Purdy, of New Haven, Conn., U.S.A., is a young man who was partly trained as a clinical psychologist and also worked for a time as an engineer. He found that he had an exceptional talent for oil painting, and decided to devote himself to this professionally. His paintings are mostly lake and woodland scenes, including some with seated human figures, and others with animals such as deer, apparently based on his experience of the forests of the Eastern United States. He paints in his studio, from memory; or rather, he constructs the paintings by intuition and imaginative use of his memory impressions, and does not use set subjects or go out into the country to copy specific scenes.

He discovered that he had a considerable colour vision defect, when he failed the colour vision tests for the U.S. Air Force, and he realized that he could not pass colour blindness tests with standard pseudo-isochromatic plates. Nevertheless, he believed that this defect, which he had been told was general for all colours, did not adversely influence his art.

The writer made contact with him as a result of reading an article in a local newspaper, and he readily co-operated in the tests and discussions, the results of which will be described here. He was glad to be given precise information about his colour vision defect as revealed by the tests, and it proved to be a specific and limited defect, not a general one as he had been told. He reported that he had a considerable degree of myopia, about 20/40, but normally wore spectacles and the myopia was adequately corrected.

DISCUSSIONS WITH THE ARTIST AND HIS WIFE

Mrs Purdy said that her husband had some difficulties with pastel shades, and that this was noticeable in particular when he tried to pick out a pastel green among blues and purples on a silk jacket which she was wearing at the time.

He called a sky-blue colour 'lilac', and, when asked to point out a 'red' colour in one of his paintings, he indicated a purple area. He expressed his willingness to mix a colour which he would call 'neutral green', and when he did this it proved to be an olive shade. He said he was particularly fond of colours called 'reddish greens', and mixed red and green to form such a colour. It was, of course, a brownish olive to the normal viewer. The use of the term 'reddish-green' is peculiar to the deuteranomalous (Pickford, 1951, pp. 99–100).

---

* A paper read to the British Psychological Society, Leicester Meeting, 3 April 1964.

When discussing his paintings Mr Purdy showed twenty or thirty of them and commented freely on his ideas about his use of colours in them. He said there had been three phases in the development of colour in his art. His early work was in dull colours. In the middle phase he followed the kind of colour scheme characteristic of the Barbizon painters, with a predominance of greys and browns. In the latest phase he had come to exploit brilliant colours in juxtaposition. He said that any 'transparent' colours would go together, and that he was always aiming at 'transparency' of effect. He asked whether brilliant colours were not very tiring to look at, producing nervous exhaustion, and said that he found painting with brilliant colours both exciting and fatiguing. He had developed the technique for using brilliant colours because he found that prospective buyers liked the paintings done in this way.

Mr Purdy explained that he worked intuitively and was a quick painter. It was apparent that he tends to make the structural parts of his paintings in blues, browns, oranges or yellows, and that he put in reds, purples and greens incidentally. He said that he found reds the most exciting of colours, but the 'reds' which he used, and to which he referred, were usually subdued or brownish. He pointed out that in his view other people, and especially prospective buyers, do not like 'red-and-green' paintings, because they are 'Christmassy', like paintings of robins and holly. His own greatest preference is for blueish greens and brownish colours, but he said that every painter would have his own peculiar tastes. It was evident that his own taste was at least partly determined by his colour vision defect.

He said he liked paintings to be restful, and asked several times whether a person with normal colour vision would find the red and green contrasts in some of his paintings offensively bright. He preferred, when using green, to use what he called a 'cool' green. This was an olive colour to normal vision. He called viridian a 'pure' green. To normal people it is on the blueish side of the green they see as most pure or uniquely green.

Mr Purdy said that he found it dazzling when small irregular areas of red and green are combined. An important point was that he said he was quite unable to paint delicate shading or shadow effects which involved subtle changes of colour. In shading and shadows he could use only changes of tone or light and dark. To illustrate this he compared an early painting of his own with an Italian picture hanging in his studio.

COLOUR-VISION TESTS

With the kind permission of Dr Jo Ann S. Kinney, and the support and interest of Dr F. L. Dimmick, Mr Purdy was tested for colour blindness in the U.S. Submarine Base, Medical Research Laboratory, New London, U.S.A., by Mr Kevin Laxar. Dr Kinney is Head of the Vision Branch of the Laboratory. The standard set of tests was employed, and in addition the Nagel anomaloscope. The tests and results were as follows:

(1) The special selection by the colour vision experts at the U.S. Submarine Base Medical Research Laboratory from the American Optical Company's set of pseudo-isochromatic plates.

Fourteen plates were used, and Mr Purdy made errors in all of them.

(2) *The double wedge hand anomaloscope.* This anomaloscope enables the subject to adjust a mixture of red and green lights to match a standard yellow. It is used to differentiate between protans, who choose more red, and deutans, who choose more green than the normal. The present subject was a deutan on this test, because he chose more green.

(3) *The Farnsworth dichotomous*-15 *test.* In this test the subject arranges fifteen coloured caps in order of nearness of hue, starting from a purple standard and going round the colour circle. This subject made no errors on this test.

The dichotomous test is designed to pass normals, mild and moderate colour defectives and to fail severe defectives and dichromats. If the subject fails the test, his type of defect (deutan, protan, tritan) is indicated by his arrangement of the buttons. Hence the fact that Mr Purdy made no errors on the test indicates only that he has no severe defect; compared with the results of the rest of the battery of tests, it becomes another indication that his colour vision is anomalous but his colour discrimination is good.

(4) *The Farnsworth lantern test.* In this lantern test two small spots of light are displayed, and the subject is asked to name their colours. Red, green and white are used, in three series of nine pairs, but in irregular order.

The subject made errors in sixteen out of the twenty-seven test items. The most frequent error was to say 'red' for white, and the next in frequency was 'white' for green. 'Red' was given for green four times, and 'green' for white three times. 'Green' or 'white' was never given for red. Since the spots are in pairs there seem to be simultaneous contrast effects in this test. White presented with green tends to look 'red', and white with red to look 'green'.

The Farnsworth lantern uses variations in the luminance of a single colour to trick the colour defective subject since he may often, in everyday life, supply the appropriate colour name on the basis of a brightness rather than a colour clue. Mr Purdy's errors are probably more a result of these brightness differences rather than colour contrast but they were typical of deuteranomalous subjects.

(5) *The Nagel anomaloscope.* This test, as used in the Medical Research Laboratory, enables the tester to vary a mixture of spectral lights (lithium red and mercury green), to match a spectral light of sodium yellow and the brightness of the yellow standard can be varied to match that of the mixture for the subject in question. The tester carefully determines the limits of the subject's range of matches, and the range of brightness required by him. This is the best critical test for colour vision defects, and shows the subject's type and degree of defect precisely. The results for Mr Purdy are shown in Table 1, and may be compared with those for normal subjects on the same individual instrument, as shown in Table 2.

From the Nagal anomaloscope result it is clear that Mr Purdy is a simple deuteranomalous (DA) subject, because, (1) he has a very marked deviation to the green side, coupled with (2) a small or moderate increase in matching range, as compared with the normal result, but, (3) his matching range does not extend so far as to include the normal mid-matching point, and, (4) his brightness matches are normal. This is decisive, and the other test results accord with it, judging by the writer's experience of colour defective subjects. [(1) On the pseudo-isochromatic test he failed outright. (2) On the double-wedge anomaloscope he was a deutan. (3) On the

Farnsworth dichotomous-15 test he made no errors. This is an indication that his colour vision is anomalous but his colour discrimination is good. (4) On the Farnsworth lantern he tended to call white 'red' or 'green', green 'red', and white 'green', but did not call red 'green' or white 'red'.]

Table 1. *Nagel anomaloscope* (*Mr R. Purdy*)

	Limits chosen	Mid-point and matching range
For colour	19–26	22·5 ± 4
For brightness	13–15	14 ± 1

Table 2. *Nagel anomaloscope* (*normal result*)

	Limits chosen	Mid-point and matching range
For colour	38–42	40 ± 2
For brightness	14–15	14·5 ± 0·5

DISCUSSION

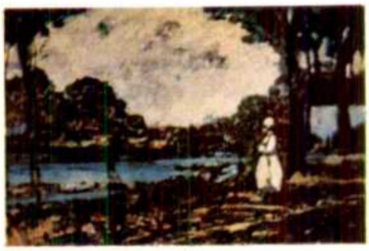
Simple deuteranomalous subjects, as is well known, have good or fairly good colour discrimination, but deviate widely from the normal in their choice of a match in the Rayleigh Equation. Mr Purdy had a matching range, which indicates colour discriminations, only twice the normal, whereas in severely red-green defective subjects it might be ten or more times larger. However, his deviation was to the green side of the equation, and was 18 units on the test, which is very large, and indicates that all colours which are greenish to the normal person tend to be yellowish to him, and that he finds his pure green among blue-green shades. Consequently he also tends to see yellows as more orange, and oranges as more reddish in colour than does the normal person.

The writer has reported on the colour vision of a woman art student who was deuteranomalous (Pickford, 1951, pp. 197–8). Her defect did not appear to have been a great disadvantage in her work, but there was no decisive evidence of this. The reason why deuteranomaly of the simple type (DA), even when it causes the person to fail in strict colour blindness tests, may not be a marked disadvantage to artists, is that the DA subjects often have fairly good colour discrimination, as already mentioned, although the colour matches they make will be unusual. Deuteranopia, extreme deuteranomaly and the corresponding protan defects would be more serious, because in them colour discrimination is more severely affected.

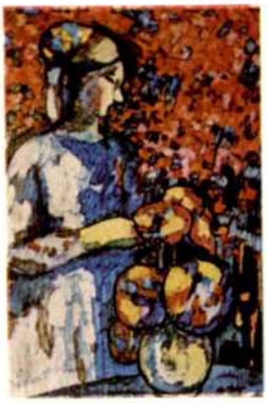
An account of Roaf's painting test for colour blindness, and a coloured reproduction of one of her subjects' pictures, are given by Collins (1925). This subject appears to have been a protan.

A deuteranope (D) copied a painting and made many errors such as: yellowish red edges of flames—bright lime green; dull red flower pot—green; olive green leaves—blue green; pink flowers—green, and so on (Pickford, 1951, p. 233). Another man who was extreme deuteranomalous (EDA) copied the pictures in Van Lennep's 'Four-Picture Test' in water colours, giving people green faces and making other mistakes such as blue green instead of grass green. A group of sixteen art students,

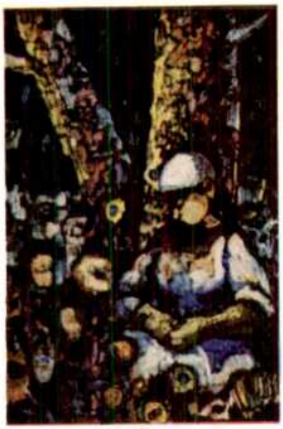
British Journal of Psychology, Vol. 55, No. 4



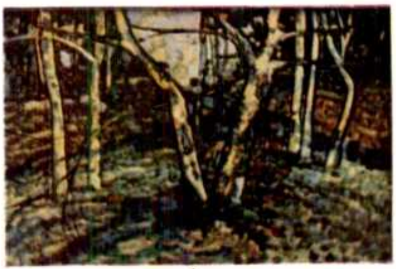
1



2



3



4

R. W. PICKFORD

(*Facing p.* 473)

thirteen of whom were women, did not differ statistically from the normal population tested as a whole, except that on a lenient statistical standard ($P < 0{\cdot}05$) they were a little more sensitive to red and green (Pickford, 1951, pp. 197–8).

Riddell (1949) has suggested that colour defective men with artistic talent may tend to be diverted into engraving. Trevor-Roper (1959) raised the question whether Constable may have been red-green defective, perhaps a protan, because of the mainly brownish colourings in his paintings, and the blue-greens which predominate in a minority of them. He mentions other painters who may have been colour defective, such as Whistler, Carrière, and Grottger, the Polish master of pencil and charcoal drawing. Some artists, he says, would avoid the use of colour, like Whistler if he were a defective, while others might use colours boldly. He points out that Strebel claimed in 1933 that the artist Léger (1881–1955) must be colour blind, owing to the blue and yellow colours which predominate in his work, and proved himself to be correct by testing the artist. He was one who used colours boldly (Trevor-Roper, 1959, fig. 33), and Mr Purdy, the artist discussed in the present paper, is another who has tended to over-compensate for his colour vision deficiency in this way, although he prefers subdued colours himself.

NOTES ON THE COLOUR PLATE

The colour plate, introduced with Mr Purdy's permission, and with financial aid from the Carnegie Trust for the Universities of Scotland, illustrates four of his paintings.

These illustrations are very good from the point of view both of colour and of detail, although they were made from the writer's colour transparencies, taken with the artist's help in his garden.

No. 1 shows the kind of colouring he prefers.

No. 3 shows his most usual way of construction with blues, yellows and browns or black.

No. 2 shows an unusual introduction of red freely, against a blue, white and yellow framework.

No. 4 shows an equally unusual introduction of green, against a framework of yellow, brown, blue and black.

Red and green are almost never used as important contrasting colours in the same painting, but yellow and blue are used in this way constantly.

POSTSCRIPT

Since writing this paper the author has received communications from Prof. George Humphrey (Cambridge) and Mr J. R. Morrison (Aberystwyth).

Prof. Humphrey writes that he has a somewhat 'modernistic' painting called 'Winter Evening with Jesus and Mary'. It has no red or green, but is exceedingly 'life-like'. Those who have never been in Canada (the scene) can hardly believe it. The artist, Thoreau MacDonald, told him that he had no red or green vision, and always had to have his father, a very good artist, put in these colours when necessary, and pass judgement on them in general.

Mr Morrison draws attention to 'Stroudley's Improved Engine Green, an unusual locomotive paint colour adopted on the Highland Railway, and on the

London, Brighton and South Coast Railway, by William Stroudley when he was Locomotive Superintendent of these companies. It was a cool shade of golden ochre'. This statement is from Ernest F. Carter, *The Railway Encyclopedia* (Harold Starke, Ltd., 1963), p. 313. In addition, Cuthbert H. Ellis, in *The North British Railway* (Ian Allen, 1955), p. 99, says, 'I suggest that...Stroudley's...eyes were somewhat unreceptive of yellow tones. *I suggest that Stroudley thought his yellow engines really were green*'. (Original italics.) 'His eyes were myopic and possibly had other defects.' 'Ellis paints himself', says Mr Morrison, 'and most of his books are illustrated by colour reproductions of his own pictures. He appears, however, to believe that there may be an association between myopia and colour-defect.'

Mr Morrison indicates that further references to Stroudley's improved engine green, and descriptions of it, occur in other books by Ellis, namely, *The Trains we Loved* (Allen and Unwin, 1947), pp. 23–4, *Some Classic Locomotives*, (Allen and Unwin, 1949), p. 85; *The Beauty of the Old Trains*, (Allen and Unwin, 1952), pp. 68–9; *Twenty Locomotive Men* (Ian Allen, 1958), p. 117; and *The London, Brighton and South Coast Railway* (Ian Allen, 1960), pp. 105–6.

Mr Morrison presented the writer with a coloured post card of the locomotive L.B.S.C., Class 'B1', 0-4-2, no. 175, 'Hayling', designed by W. Stroudley and built at Brighton in 1890. It is painted in the 'improved engine green', and is a golden yellow colour, with crimson lines and black funnel. From the evidence presented, it is almost certain that Stroudley was a dichromat and a deuteranope. Myopia is not the cause of colour vision defect. If he had been a protanope he would have confused crimson and black and not included both as distinct aspects of the same colour scheme. The postcard is published by Messrs Ian Allen, Ltd., Terminal House, Shepperton, Middlesex.

REFERENCES

COLLINS, M. (1925). *Colour Blindness.* London: Kegan Paul.

PICKFORD, R. W. (1951). *Individual Differences in Colour Vision.* London: Routledge and Kegan Paul.

RIDDELL, W. J. B. *et al.* (1949). Discussion on colour vision in industry. *Proc. R. Soc. Med.* (*Ophthalmol.*), **42**, 145–50.

TREVOR-ROPER, P. D. (1959). The influence of eye disease on pictorial art. *Proc. R. Soc. Med. Ophthalmol.*), **52**, 721–44.

(*Manuscript received* 26 *March* 1964)

Brit. J. Psychol. (1964), 55, 4, *pp.* 475–476
Printed in Great Britain

SEASON OF BIRTH, INTELLIGENCE AND PERSONALITY MEASURES

BY ANN D. M. DAVIES

Medical Research Council Unit for Research on Occupational Aspects of Ageing, at the University of Liverpool

In recent articles Orme (1962, 1963) reports significant differences in level of intelligence associated with season of birth. In a group of intellectually subnormal subjects he found that those born in the summer and autumn months had significantly higher 'I.Q.' scores than those born in the winter and spring months. The 'I.Q.' scores were derived by Orme (1961) from scores on Raven's Coloured Progressive Matrices test. Orme suggests that the temperature of the months preceding birth may influence subsequent adult intelligence. The differences he finds are in the same direction as those found by Fitt (1941), but Orme (1962) considers that these early studies used techniques 'which would now be considered outmoded'.

If the differences in non-verbal intelligence reported by Orme are found to be reliable they could affect psychological research profoundly. They would make it as necessary to control for season of birth as for the variables—age, sex and social class.

The present study was organized to see whether Orme's results could be confirmed for a population of normal adults. The biographical data sheets and test scores of the Medical Research Council Merseyside voluntary research panel were analysed, and the following variables examined:

1. Non-verbal intelligence (Progressive Matrices, 1938).
2. Vocabulary level (Mill Hill Synonyms Senior B).
3. Emotional maladjustment (Heron, 1956).
4. Sociability (Heron, 1956).

Approximately 300 men from the Merseyside panel were used as subjects, fifty men in each of the six age decades 20–29 to 70–79. The panel (Heron, 1962) is made up of people drawn from a wide variety of occupations and intelligence levels. Within each age-group there is a close approximation to the national distribution for socio-economic classification. Thus the panel as a whole is a fairly representative sample of the general population. The panel was divided into those born in the summer and autumn months (May–October) and those born in the winter and spring months (November–April). The numbers in these groups were 148 and 151 respectively. As some of the test scores to be studied are known to vary with age, each age decade was examined separately on one variable at a time. The median score for each age group was found and the group split into those with scores falling on or above the median and those whose scores fell below the median. A 2×2 χ^2 test was calculated to ascertain whether significant differences existed between the summer–autumn group and the winter–spring group.

No value of χ^2 was statistically significant for any variable or for any age-group. When the age-groups were combined, there were still no significant differences.

In order to ensure that significant differences were not obscured by combining summer with autumn and winter with spring births, the data were re-analysed combining spring with summer and autumn with winter. Once again no significant differences were found between the test scores of the panel split in this way, nor were significant differences found on non-verbal intelligence when the medians quoted in the current Guide to the Standard Progressive Matrices (Raven, 1938) were used as the cutting point.

Orme (1962) states: 'That different birth months can exert such an effect, even within a subnormal group, suggests that this effect is marked and persistent.' The present study shows that in one population, at least, season of birth has no effect on two tests of intelligence and two measures of personality. Whether the differences found by Orme are a peculiarity of subnormal populations remains to be seen.

REFERENCES

FITT, A. B. (1941). *Seasonal Influence on Growth, Function and Inheritance.* London: Oxford University Press.

HERON, A. (1956). A two-part personality measure for use as a research criterion. *Brit. J. Psychol.* **47**, 243–51.

HERON, A. (1962). The Liverpool functional age project: preliminary communication on population and methodology. *Experientia*, **18**, 470.

ORME, J. E. (1961). The Coloured Progressive Matrices as a measure of intellectual subnormality. *Brit. J. Med. Psychol.* **34**, 291–2.

ORME, J. E. (1962). Intelligence and season of birth. *Brit. J. Med. Psychol.* **35**, 233–4.

ORME, J. E. (1963). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.* **54**, 273–6.

RAVEN, J. C. (1938). *Guide to the Standard Progressive Matrices.* London: Lewis.

RAVEN, J. C. (1956). *Guide to the Coloured Progressive Matrices.* London: Lewis.

(*Manuscript received* 21 *October* 1963)

Brit. J. Psychol. (1964), 55, 4, *pp.* 477–484
Printed in Great Britain

BAUDOUIN ON JUNG

By CYRIL BURT

Owing to his long and intimate friendship with Jung and his family Professor Baudouin has been able to throw a revealing light on many of the obscurer features in Jung's psychological theories. Their development, he believes, was largely the result of Jung's own personal problems and of certain emotional incidents in Jung's private life. In his survey of Jung's publications, and in tracing the progress of his ideas, Baudouin, unlike most previous commentators and reviewers, endeavours to bring out the special bearing of Jung's teaching, not so much on the problems of psychiatry and abnormal psychology, but on those of general psychology; and he has thus drawn attention to many neglected aspects of Jung's work.

Charles Baudouin was born at Nancy and trained in Paris under Bergson. But the last forty years of his life were spent at Geneva, where he became Professor at the University, and where he died last August at the age of 70. His earlier books on *The Mind of the Child*, *The Psychoanalysis of Art*, and, most of all, *Suggestion and Auto-suggestion*, aroused keen interest over here; but his later works—mainly psychological studies of French and Swiss authors—have naturally attracted less attention. As a poet and a versatile essayist he writes on the most obscure subjects in a style that is at once lucid and readable; and in his latest book, published posthumously, he succeeds far better than Jung himself has ever done in presenting Jung's highly original speculations and conclusions in an intelligible and systematic shape.

As a close friend of Jung for over thirty years, he came to know him far more intimately than any other psychologist; and one of his chief aims, he tells us, has been to 'study, as it were, the psychology of Jung's own views upon psychology'. From his wide knowledge of Swiss literature and Swiss religious and philosophical thought, he is able to show how much Jung owed to the cultural heritage of that country—to the early Swiss humanists, to the genetic psychology of Claparède, to the psychiatric and parapsychological studies of Flournoy, and particularly to the mystical epics of Carl Spitteler, the Nobel prize winner. The unfamiliar facts which he has thus brought together throw a new and revealing light on much in Jung's books and papers that readers in other countries have found obscure and even bizarre.

Professor Baudouin begins with a biographical introduction describing the main stages of Jung's intellectual development, and illuminates them with vivid personal reminiscences. He notes that many of Jung's interests and characteristics are traceable in those of his forefathers. How far their reappearance was due to heredity, tradition, or deliberate emulation, he leaves us to guess. Jung himself, in discussing 'the fateful links between myself and my ancestors', declares that 'it seemed as though there had been a personal *karma* within the family, passed on from parents to offspring'. The first Carl Jung, who died in 1645, was a doctor of medicine at Mainz, at that time a centre of alchemy. Carl Gustav Jung, the psychologist's grandfather and namesake, was reputed to be an illegitimate son of Goethe, and was himself a poet and freethinker; in the family's coat of arms, instead of the traditional phoenix (emblem of youth and a heraldic pun on the family name) he substituted a set of Rosicrucian symbols. Exiled from Germany because of his revolutionary ideas, he settled in Basel, where he occupied the chair of surgery. His portrait still hangs in the University Hall.

Jung's father was a pastor in the border hamlet of Kesswil. We are given passing glimpses of little Carl—first as *petit paysan* romping round the falls and castles of the Rhine, then as a precocious pupil at the village school, secretly cherishing intellectual ambitions of his own, and later as 'a citizen of no mean city'—the ancient town where Paracelsus had lectured, Holbein and Erasmus made their home, Nietzsche held a chair of classical philology, and which in Jung's early youth could boast a Spitteler, a Boecklin, and a Burckhardt. In such surroundings, as his biographer observes, Jung's own imaginative flights 'seem almost a natural growth'.

Hitherto most of the studies of Jung's prolific output, including those published in this country, have concentrated chiefly on the medical and psychotherapeutic aspects of his work. Baudouin's aim is to examine its significance for general psychology. He divides his survey into four main parts. In the first he describes what he considers to be Jung's '*idées directrices*' and their sources in other continental writers—Janet, James, Bergson, Nietzsche, von Hartmann, Bleuler, and Freud. The second part deals more fully with Jung's typology and the hypotheses on which it is based—the doctrines of a general psychic energy ('libido' or 'horme') and the 'bipolar tension of opposites'—with the formation of 'emotional complexes' as an important resultant. Part III, entitled 'the reality of mind' (Jung's own phrase), analyses his conception of the self, which, following James, he regarded as comprising 'a number of different selves, one within another'—the *persona* (James's 'social self'), the 'me' (the 'empirical or conscious self'), the *anima*, the *animus*, and the *umbra* (constituents of the 'unconscious self'), and the 'pure ego' (analogous to Kant's noumenal self, except that it is not entirely unknowable). This leads to an examination of two related concepts to which Jung attaches special importance, but which he never seems able to make perfectly clear—the 'collective unconscious' and the 'archetypes'. Part IV, the longest in the book, reviews the amazing mass of work published during what Jung himself describes as 'the second and most fruitful half of my long life'—his various excursions into Eastern and Western mythology, his abstruse disquisitions on the fanciful theories of medieval alchemists and astrologers, the mystical speculations of Taoist and Hindu philosophers, and finally his own views on metaphysics and religion. The book ends with a brief comparison of the rival schools of psychoanalysis regarded as attempts to develop a practical psychology of motivation, and a survey of Jung's influence on later continental writers—scientific, theological, and philosophical, as well as psychological.

There remains, however, one puzzling question, which has often been raised and which Baudouin confesses himself unable to answer—why did Jung abandon his investigations into the nature and treatment of mental disorders to become absorbed in out-of-the-way topics like alchemy, oriental mysticism, and the prophecies of Nostradamus? A partial explanation has now been given by Jung himself in his recent autobiography (Jung, 1963), which unfortunately appeared too late for Baudouin to make full use of. There Jung reveals how he himself became the subject of mystical experiences, and discovered that he was apparently endowed with paranormal powers—telepathy, precognition, and even psychokinesis.* And it was, so he tells us, partly to gain an understanding of his own inner life that he felt driven to search for historical parallels. These he at length discovered in the writings of the alchemical school—especially those of his own compatriot, Paracelsus. Here, he believes, 'is the bridge that has led from early gnosticism to the modern psychology of the unconscious'.†

Most British commentators on Jung's psychology tend either to ignore these digressions into the occult sciences or to dismiss them as an old man's aberrations. Prof. Peters, for example, finds 'much of Jung's later work, e.g. what he wrote on "The Psychology of Religion," so mysterious as to be almost undiscussable' (Peters, 1953, p. 695); and Dr Glover dismisses it all as 'a mish-mash of oriental philosophy with a bowdlerized psychobiology' (Glover, 1950, p. 134). On the other hand, contemporary chemists, like J. Holmyard and H. M. Leicester, speak of his erudite studies as 'an illuminating contribution to scientific history'.

However, let us briefly glance at Jung's own account. The alchemists, as he points out, were almost the only experimentalists of their day, and in their fashion they even 'sought to apply the experimental method to psychical and psychiatric phenomena'. What is more, they were the first to treat religious and mystical experience, not as a supernatural revelation, nor yet as a purely speculative philosophy of the cosmos, but as something that could be described and discussed scientifically—rather in the spirit of the modern parapsychologist. 'Above all', he says, 'their experiences were my experiences; their situation that of my patients.' The majority of those who came to him were people who 'could not accept the orthodox faith, and the conflicts thus arising were among the most frequent and the most powerful causes of their neurotic disturbances'. A similar inability to accept the orthodox faith was characteristic of the renaissance alchemists. They too were intellectual renegades. (It is a quandary vividly depicted in Browning's poem, *Paracelsus*.) Freud told his patients they should be thankful they were rid of an infantile illusion. Jung, as a result of his own mystical visions, became convinced that 'religion—or rather the elusive basis of all religions—was not something wholly factitious'; and concluded that what he had to find for his patients, and indeed for sceptical humanists in the world at large, was 'an equivalent over-belief'. At the same time 'it was necessary to avoid a formulation which would remind them too closely of what they had rejected'. The basis for this 'equivalent' he found in the modified Neo-Platonism of the Hermetic writers as expressed in the vivid symbolism of the medieval alchemists, who 'in their role of physicians dealt with diseases of mind as well as body'.

---

* Much of the book—nearly all that is new in it—consists of vivid descriptions of Jung's manifold experiences and of the influence they exercised on his theories and his writings: (I have discussed them in detail elsewhere (Burt, 1963), and so need not enlarge on them here). It should be noted that, except for one striking incident—oddly enough attested by Freud in an extant letter—hardly any of Jung's claims are corroborated by extracts from diaries, letters, or reports of other witnesses written at the time; some are demonstrably inaccurate. The critical reader cannot help wondering how much has been embellished by the elaborations of a retrospective memory or a lively imagination, or what allowance must be made for the fact that most of the chapters were not written by Jung himself, but compiled by Miss Jaffé on the basis of belated interviews when Jung was well over eighty. Jung indeed describes them as his own 'personal myth': 'whether or not the stories are "true"', he adds, 'is not the problem'.

† The student should note that at times Jung goes badly astray: e.g. he explains that Poimandres was 'a pagan gnostic'. As Betsy Prig might have said, there never was such a person. The title appears in several books of the *Corpus Hermeticum* (a Greek fourteenth-century MS. containing the remnants of an extensive literature produced in Egypt during the first two or three centuries of our era). In ancient Egyptian the word would mean 'knowledge of Ra' (as the embodiment of the Good); the Hellenized form may have been intended to suggest the 'shepherd of man'. Alchemy was not, as Jung supposes, an outgrowth of gnosticism, but a parallel development. The secret doctrines of 'Al-Khem' (i.e. 'the Egyptian science') go back to the odd jumble of Egyptian technology, Babylonian astrology, and Greek Neo-Platonism, overlaid with Jewish and Persian religious ideas, which developed in Alexandria during the Hellenistic period, and found its most notable expression in the Hermetic writings. It fascinated the Arabs of the eighth and ninth centuries after they conquered Alexandria; and (as the alchemist's vocabulary indicates) both the theoretical speculations and the practical techniques came to Europe through Arabian channels. C. H. Dodds and others have shown how these Hermetic doctrines influenced, not only gnosticism and alchemy, but also the development of Christian philosophy from the Pauline and Johannine writings onwards and even scholastic psychology; but here the influence was collateral rather than direct.

As adequate information was hardly available at the time he wrote, Jung is still more inexact in his interpretations of Chinese psychology. The student who is interested in this part of Jung's work would do well to turn to Prof. Fung Yu Lan's *Short History of Chinese Philosophy* (1960), and especially to Wing Tsit Chan's *Source Book in Chinese Philosophy* (1964).

'Alchemy', says Holmyard, the leading authority on the history of the subject, 'was of a twofold nature: exoteric alchemy was concerned with the "philosopher's stone", which had the power of transmuting "base" metals into "noble"; esoteric or mystical alchemy was concerned with the *elixir vitae*, and developed into a devotional system whereby the mundane transmutation became symbolic of the transformation of sinful man into a perfect being' (Holmyard, 1957, pp. 13f.)[*]. And it is the 'esoteric' aspects of alchemy that interest Jung. But why need it be 'esoteric'? 'Why', asks Father Maretti in a letter to Baudouin, 'should either alchemist or medical psychologist wish to express their doctrines in these occult and enigmatic symbols? Even if they feel that some kind of concrete analogy will be helpful to their disciples or their patients, why should they prefer the fanciful and unfamiliar process of transubstantiating lead into gold to the far more natural symbolism that talks of bread and wine as representing the heavenly nutriment which satisfies our spiritual hunger and thirst? Or why substitute the highly specialized techniques of burning impurities in the chemist's furnace for the time-honoured symbolism of washing them away in the blood of the lamb? Is it a case of *omne ignotum pro magnifico* or just *ignotum per ignotius*?'

There are, I think, several plausible answers. First of all, as Jung points out, the stereotyped Judaeo-Christian imagery had become associated with antiquated dogmas which neither the alchemist of the early Renaissance nor the patient in the modern consulting-room was willing to accept; as a psychotherapist Jung himself preferred a symbolism which was more flexible and less likely to be taken literally.

---

[*] Compare Sherwood Taylor, *The Alchemists* (1949). It is unfortunate that literary references to alchemists, from Chaucer and Ben Jonson onwards, have encouraged the common notion that they were nothing but quacks pandering to vulgar credulity and greed. Paracelsus's special brand of medieval Latin is singularly difficult to follow. The English reader will find the cardinal doctrines set forth with admirable clarity by Robert Fludd, son of Queen Elizabeth's Treasurer in the Low Countries. At once an experimentalist, an Oxford scholar, a disciple of Paracelsus, a Rosicrucian, and a Fellow of the Royal College of Physicians, he retains the systematic logical style of the scholastic disputation, using formal syllogistic proofs. Hermes Trismegistus is quoted as an authority of equal rank with Aristotle, Plato and Pythagoras. Fludd's panpsychic cosmology, and his free use of symbolism, was repeatedly attacked by Kepler; but, replies Fludd, 'I use these symbolic models (*figurae hieroglyphicae*), not because I delight in pictures, but because the *arcana* of science are often better conveyed by this convenient device than by current or freshly coined abstractions or your own numerical *gematria*' (*Replicatio in Apologiam*, Frankfurt, 1622: cf. also Thomas De Quincey's delightful study, *Rosicrucians and Freemasons*, 1855). Kepler's hypotheses were based on much the same 'occult' sources as those of Fludd. But Kepler changed the whole course of science by introducing mathematical calculations based on observed measurements. The doctrines of the astrologer and the alchemist sank into oblivion, or survived only as a popular superstition, and their psychological implications were forgotten until Jung brought them to light.

But secondly, the symbols and the symbolic language of the Judaeo-Christian religion were suited to a primitive sheep-rearing and agricultural community, whereas the language and the thought-models of sixteenth-century iatro-chemists seemed far more appropriate to a civilization that was making an increasing use of chemical, metallurgical, and other scientific processes. However, with the alchemists, I suspect, there were two further reasons, which Jung seems to overlook. The chemical and metallurgical processes of the professional alchemists were trade-secrets, for the products of which a handsome price was exacted; hence the technical manuals written or printed for the use of authorized practitioners were couched in a cryptic language based on psychophysical analogies. On the other hand, the advanced philosophical theories of these scientific pioneers were universally condemned as dangerous heresies by Catholics and Protestants alike: they had therefore to be wrapped up in a converse astro-chemical symbolism, whose real significance would be intelligible only to members of some secret society—the Rosicrucians, the Freemasons, or the like—which was powerful enough to lend its protection. And no doubt, as Father Maretti suggests, the game of mystification was amusing in itself. Certainly today anyone who, tired of solving his daily cross-word puzzles, cares to spend his leisure hours trying to disentangle the hidden implications of, say, Mylius's *Philosophia Reformata* or the ostensibly anonymous *Ordinall of Alchimy* (the full name of the writer is in fact revealed by the initial letters of the chapters) will enjoy, on a somewhat erudite level, much of the thrill of unravelling an ingenious detective story or some historical mystery, like the identity of 'Mr W. H.' or 'The Man in the Iron Mask'.

The speculations of the alchemists, says the poet Victor Hugo (aptly quoted by Prof. Baudouin), were 'les rêves de la médecine, quelquefois ses cauchemars'; and Jung, boldly following a suggestion of the young psychoanalyst Silberer, began by analysing their traditional symbols and obscurer sayings much as a Freudian would analyse a sequence of dreams. The strictly empirical psychologist will doubtless be tempted to say, with Bottom the weaver: 'A man is but an ass if he go about to expound this dream'. Nevertheless, Jung's method unquestionably has proved both fruitful and successful.

Underlying all the teaching of the 'esoteric school' of alchemy Jung claims to find 'a profound philosophy which is nothing if not practical, and therefore a religion rather than a mere metaphysic'. Himself the son of a pastor, he was throughout his life intensely interested in the psychology of religion. He notes, as others had done, that the elaborate *opus* of the alchemists displays striking parallels with the rites of the higher types of religion—with the initiation ceremonies of the Orphic mysteries, for example, as well as with the Christian sacrifice of the mass. And he argues that these resemblances are due, not to conscious borrowing or parodying (as has been generally averred), but to the fact that the essential elements common to them all formed 'the natural expression of wishful tendencies and inarticulate feelings shared by the whole human race'—in short, of the 'primaeval archetypes'. Aesthetic, emotional, and mystical experiences, he insists, cannot be adequately conveyed by words; they have to be suggested by semi-poetical or semi-conventional symbols, images, and myths. Only thus can one utter the unutterable.

Jung is often blamed for inconsistency. But, as with every original thinker, his novel ideas were naturally altered or reshaped as they evolved. In the development of his religious views we may, says Baudouin, distinguish three main phases. At the outset he was content to accept a materialistic and positivist attitude similar to that of Freud and of most continental physicians at the turn of the century. After the break with Freud he adopted a deliberately non-committal standpoint, and on philosophic matters sought to preserve 'un agnosticisme absolu'. 'On a prétendu que j'étais un philosophe', he observed to Baudouin, 'en réalité je suis un empirique.' Nevertheless, throughout his later writings he was, as he acknowledges in his writings, deeply influenced by James's book on *The Varieties of Religious Experience*. Finally, in his third 'post-therapeutic' phase he frankly announced 'un au delà de l'agnosticisme', vague at first, but becoming more and more explicit in regard to details as his enthusiasm increased.

These successive shifts in his theoretical outlook, so Baudouin tells us, seemed to be 'in close correspondence with the inner development of his practical life'. And quite independently Prof. Meier, one of Jung's leading disciples at Zürich, remarked that 'Jung's thoughts always appeared to have had an intimate connexion with his private life at the time'. Dr Avis Dry, who reports the comment in her admirable book on *The Psychology of Jung* (Dry, 1961), goes on to regret 'how surprisingly little is known about Jung's life and development'. The gap has now very largely been filled by the autobiography. The motto on the very first page is a quotation from Coleridge: 'He looked at his own Soul with a Telescope.' Miss Jaffé, his amanuensis and editor, tells us in her preface that Jung himself regarded it as 'his religious testament'. And certainly, as she puts it, this is 'the only place in which he speaks of his own experiences of God'. He is fond of referring to the apostle Paul; and the book is the story of his own road to Damascus.

As a thorough-going empiricist Jung at first eagerly embraced James's 'pragmatic view of truth'. Indeed, in dealing with his patients, he seems at times to have been ready to offer them a religious *placebo*, much as some physician might prescribe an innocuous tonic—not because he believed in its objective truth, but because, if his patients believe in it, it will accelerate their cure. Thus, in his earliest writings—in *The Psychology of the Unconscious*, for example—God was depicted merely as a 'projection of man's own libido', or (a little later) as a reification or personification of the 'collective unconscious'. He cites James's dictum: 'Whatever it may be on the *farther* side, the "more", with which we seem to feel connected in our religious experience, is on its *hither* side merely a subconscious continuation of our conscious life.' But, as Dr Dry remarks, 'the therapeutic value of religion must always be something of a by-product, depending on the religion being accepted as true in and for itself'. This Jung himself came gradually to recognize, partly as a result of observations on the after-histories of his patients, but still more (as we now learn from his own lips) in consequence of his mystical experiences. 'I am a splinter of the infinite Deity', he now declares, 'and my life is the story of the self-realization of the unconscious.'

However, in Baudouin's view by far the most valuable of the many contributions made by Jung to general psychology is his theory of motivation. Baudouin attempts a synthesis of Jung's theory with that of Freud; and the outcome is a structural conception of the human mind on its conative side not unlike that put forward by

McDougall. It presents, as Baudouin puts it, 'le caractère remarquable d'être *hiérarchique*' (his italics): 'that is, it depicts the human mind as consisting of three superposed layers, which roughly correspond to the main stages of racial evolution and of individual development'. In contrast to Freud's earlier and narrower definition, Jung and McDougall retain the older and broader view of unconscious processes, and insist on their importance on every mental plane; and Freud himself in his later writings came gradually to accept this suggestion along with many others put forward by Jung. (1) The lowest layer consists of the innate impulses or 'instincts'; and both Jung and McDougall maintain that such tendencies must include, not only the primary emotional drives, but also certain perceptual aspects (the immediate recognition of particular objects or situations as having an implicit meaning and importance), as well as certain conative elements (the 'implicit consummatory goals'). (2) The next level—the associative—consists of the habits, memories, and skills acquired by the individual during his life-time: once again both writers agree in emphasizing that at this stage the associations giving rise to such formations are for the most part unconscious, and that the thinking is, in Bleuler's phrase, 'autistic'. These associative systems also have an emotional aspect, varying in intensity; and, as McDougall remarks, Jung's concept of the 'complex' corresponds very closely with what he had already described under the name of 'sentiment'. Of these by far the most important is the 'ego-complex' or 'self-regarding sentiment'; and here, as we have seen, Jung's analysis is far more detailed than either that of McDougall or of Freud. (3) Finally in man, in contrast to the animals, there is yet another level—the ideational or conceptual: at this stage both the guiding ideas and the associations or relations between them tend to become explicitly conscious, though much of the mental organization is itself still carried out unconsciously. This uppermost layer is the source, not only of rational knowledge and rational conduct, but also of the higher imaginative activities and processes by which we perceive and judge the higher values.

The healthy development of the normal individual, we are told, involves a process of 'individuation', or (to use an expression common to both McDougall and Jung) an 'integration of the personality'. In his efforts to attain this 'inner harmony and stability', the child, the youth, and even the adult, generally require a certain modicum of guidance. In the past this 'spiritual assistance' has been provided by the teacher and the priest. But now, says Jung, the psychologist must step in and lend his aid. It was largely on these grounds, he explains, that he decided to turn his attention from the neurotic to the normal, and so proceeded 'de la thérapeutique à l'art de vivre'. It was his growing appreciation of 'the urgent psychological needs of the times' which, at first half unconsciously, and then somewhat reluctantly, led him to undertake 'a critical study of the psychology of social morality and of traditional religion, with a view more especially to determining their present limitations. 'This', he adds, 'the psychology of the textbooks has failed to do.' The theoretical psychologist (to quote Baudouin's comment once again) is, and should be, a scientific empiricist and nothing more; the practical psychologist must also be *un philosophe malgré lui*.

Brit. J. Psychol. (1964), 55, 4, *pp.* 485–506
Printed in Great Britain

PUBLICATIONS RECENTLY RECEIVED

Psychology. The Science of Mental Life. By George A. Miller. London: Hutchinson. 1964. Pp. xi + 388. 30*s*.

When a reviewer writes that it is not clear to what audience a book is addressed, the implication is usually that it is suitable for none. If one expresses a similar doubt regarding Dr Miller's book it is because of a conviction that it could be read with pleasure and profit by anyone. Its inclusion in 'The Hutchinson Science Library' suggests a non-specialist treatment, and the book certainly lacks the detailed tables and other characteristic features of the expensive large-format introductory texts. Perhaps for this very reason university teachers in this country have already found it invaluable for 'orientation' courses, and it would appear equally likely to appeal to the informed general reader—or, for that matter, to the academic or professional psychologist, who will find in the book much that is fresh and thought-provoking. Superficiality, in short, is the last quality one would impute to it.

In structure the book is unique. The orientation is historical, but the arrangement is in the approximate order in which psychologists developed the various branches of their science and the discussion of each topic is centred round a key figure. Six such persons are chosen, and on reflexion one agrees that the selection is absolutely right—Wundt, James, Galton, Pavlov, Freud, Binet.

It is perhaps an accident, but a happy one, that all six are of different nationality, for Dr Miller's writing is cosmopolitan, in the best sense. The style is terse, individual and, when appropriate, apophthegmatic. Dr Miller is at his best when drawing threads together; for example, in his chapter on 'Goads and Guides' he concludes: '...Focusing simply on the psychological problems of direction and control, it seems that several independent accounts of motivation from physiology, from psychoanalysis, from economics—tend to converge. The organism struggles to reduce the mismatch between its own criteria and perceived reality....'

Indeed, one of Dr Miller's main achievements is his successful presentation of Psychology as a unitary science. Another is the sheer readability of the book—a quality which should go far towards ensuring it a place among the even yet small number of psychological classics.

B. SEMEONOFF

Psychology and Social Problems. By Michael Argyle. London: Methuen. 1964. Pp. 232. 25*s*.

This is not a book written primarily for psychologists, but nerverthelees one that we should notice and support. Argyle has the courage of his convictions and has done no less than present a résumé of the findings of psychology as they are relevant to current social problems, addressed to the general public and more particularly our legislators. He goes further than a simple account, and has the temerity to suggest that it is time these form the basic arguments for social action. He says that the present astonishment and confusion at the pressure of social problems and the rate of social change can be dealt with only by means of large-scale social engineering. Existing methods which consist essentially of the committee approach are 'quite unsuited to the solution of concrete empirical questions. No matter how many eminent men of sound judgement hold a certain opinion, it is much more important to know the results of sound empirical investigation—the technicalities of which may indeed escape the committee members. ...Although our empirical knowledge is not complete, there are large areas of agreed findings'. Large areas of agreed findings are indeed covered. There are nearly 400 references.

The introduction gives a short explanation of basic terms which are necessary background in the fields of personality, motivation, and social interaction. The problems then reviewed are those concerned with agression, delinquency, mental disorder, prejudice, and human problems in industry. In each case the relative contributions of heredity, childhood experience, and sociological factors are discussed. Special attention is paid to the fundamental and generalized problem of social change as such.

Having analysed the problems, Argyle goes on to advocate a series of Processes of Social Control. Stylistically, this separation of problem and control has resulted in a divorce of subject-matter leading to a regrettable degree of repetition, over-lap, and cross-reference. However, in terms of content this is the most stimulating section of the work. Argyle operates here on a truly practical level, and puts forward concrete strategies in the fields of personality formation and modification, leadership techniques, communication, and educationtional selection which will result in more effective, harmonious and fulfilled living. For example, the achievement motive is considered. First, there are described the conditions conducive to its development in a particular child in the context of a western nuclear family. It is suggested that we present our children with an image of *visible* hard work, and that we be generally available, warm, and attractive objects of identification. Argyle then follows McClelland in saying that increasing achievement could profitably be a primary goal of under-developed countries and perhaps also of certain religious subcultures in our own societies.

The book is written with the author's characteristic, and now familiar, gusto. This does not mean that it lacks either care or sophistication. For instance, as well as presenting ideas essentially novel to a wider public, trouble has been taken to revise certain precepts that have already gained general acceptance. It is noted that the doctrine of the efficacy of permissive, almost non-leadership, is modified; as is the belief that social factors provide the most important form of satisfaction in all kinds of work, Argyle's own explosion of the Hawthorne studies being quoted here. His recent pre-occupation with the problems of introjection, identification, and socialization is clearly evident. And here again the point of view is a revised one, the advocacy of withdrawal of love as a technique in child rearing. Karen Horney, among others, laboured hard and probably successfully to persuade us of the harm that can accrue from conditional love. Now, based primarily on Sears's work, it is suggested that the formation of conscience is dependent on at least a related mode of psychological discipline. It is interesting to see that we have now aligned ourselves with Soviet pedagogues who, working from rather different premises, have always supported this method.

Broadly, Argyle's recommendations may be accepted as scientifically unobjectionable. And it may be agreed that rational social planning is essential in a modern state. Like economic planning, it increases the freedom of the individual rather than the power of the state. 'With planning, society changes in a manner determined by democratic procedures on the basis of fully informed discussion; without planning, society moves under the influence of blind, uncontrolled, social forces.' However, the specific procedures do sometimes involve political manoeuvres, the complexity of which has not always been appreciated. It is true that decrease in racial prejudice would be achieved if large numbers of coloured people were prevented from suddenly pouring into a particular area or occupation—but passing regulations to that end is not a simple matter.

HALLA BELOFF

Fields of Applied Psychology. By ANNE ANASTASI. New York and London: McGraw-Hill. 1964. Pp. xi + 621. 69*s*. 6*d*.

It is clearly stated in the introductory chapter that the main aim of this book is 'to bring together what the well-educated person needs to know about the professional activities of psychologists in business, industry, advertising and marketing, education, clinical practice, law, government and military services'. The information is directed to undergraduate and post-graduate students of psychology and other disciplines, and to professional and business people whose work brings, or should bring, them into contact with psychologists. A book directed to such a variety of potential readers may easily turn out to be of little value to any level of reader. Prof. Anastasi has avoided this danger by adopting a descriptive rather than evaluative approach by discussing the functions performed by applied psychologists but not the organizations in which these psychologists work, and by the virtual elimination of technical jargon. Short summaries at the end of each chapter enable the most hard-pressed reader to taste the fare provided, while extensive chapter reference lists direct the more serious student to important sources of detailed information.

The largest part (405 pages of text) of the book is devoted to discussion of the principal techniques and research findings of personnel, engineering, consumer, clinical and counselling psychology (Prof. Anastasi's terminology and order of presentation), while the contributions of psychologists to education, medicine and law are examined in a further 98 pages. (Appendices list Divisions of the A.P.A., American directories of psychologists, American ethical standards for psychologists and American test publishers, and show a specimen job description.) The result is a thorough and commendably unbiased account of work done by psychologists in the American community. It is easy to find out, for example, the nature of engineering psychology, or of the psychotherapeutic work performed by clinical psychologists. But a good deal of cross-reference is required before one can collect together all the book has to say about the work of prison psychologists, or of psychologists in the education services, in mental hospitals or in one of the armed forces. In other words, techniques and research and investigatory methodology common to applied psychologists are set out clearly, but much less attention is paid to the role, in its widest sense, of the psychologist in his employing organization.

How useful will this book be to the British student of psychology? The undergraduate student should gain a good idea of the sorts of work carried out by applied psychologists, but will need supplementary information about the organization of professional psychologists in the British Isles and about the working situations into which he may be plunged. To the postgraduate student the most important contribution may well be the emphasis on the large amount of ground common to clinical, educational, delinquency and occupational specialists. But if, as may be expected, he is largely concerned with learning his trade in Britain, he will wish to take into account also the structure of applied psychology in Britain. He may be taken aback by the overwhelmingly American lists of references—out of 1779 slightly overlapping 'mentions', less than fifty are by British authors living in Britain at the time of publication, and at least ten of these authors would make little claim to be called psychologists. (It should of course be borne in mind that the book is written for an American readership.)

Despite the points noted, the book is one which should be made available to all psychology students. It should be helpful to teachers of psychology, including extra-mural teachers, whose courses are concerned with activities of psychologists in the community. It is not intended to be, and should not be used as, a text-book of applied psychology.

DERMOT STRAKER

Readings in Psychology. Edited by JOHN COHEN. London: Allen and Unwin. 1964. Pp. 414. *52s. 6d.*

Considering the wide, deep coverage of these 'readings' this book appears to be remarkably relevant for postgraduate students, university teachers and research workers in various branches of psychology. The editor begins by presenting a personal 'Perspective for Psychology', followed by more than twenty chapters contributed by experts in particular branches of the field.

In the first chapter, 'Experiment in Psychology—a Centenary and an Outlook', by R. C. Oldfield, we are taken on an historical tour. Starting with Fechner's 'Elements of Psychophysics' Oldfield leads us through Wundt's 'brass instrument' experimentation on sensation; then, through the Ebbinghaus experiments on memory of nonsense syllables to Bartlett's demonstrations of the reconstructive character of memory. Oldfield next discusses two schools—Gestalt and Behaviourism—which were functioning well between the 1910's and 1940's. The ablation experiments of Franz (1902) and of Lashley (1932) are discussed in relation to their effects on learning. Finally, Oldfield advises keeping psychology within the framework of biological ideas; the next centenary should then prove one of 'exciting progress'.

In the next article, Sir Cyril Burt presents us with a vivid picture of Francis Galton's contributions in the field of individual differences and the development of tests for measuring and statistics for analysing these differences.

The next three chapters deal with theoretical problems. First, G. P. Meredith discusses the role of material or formal models; appropriate model building is held to favour research; the present state of knowledge in neurology suggests the use of formal instead of neural models.

In chapter 5 we are given a fairly thorough résumé of Jean Piaget's work on the course of cognitive development from birth to 15 years approximately, by the author Barbel Inhelder, Professor of Child Psychology, University of Geneva, who has been working with Piaget for more than twenty years. Prof. Mace presents the last of the three chapters which deal with theoretical issues. Homeostasis, Needs and Values can be interpreted to cover more than overt goal-directed behaviour; perceptual and thought activities might also be accounted for.

In a chapter on 'Physiological and Experimental Psychology', Prof. Zangwill describes the results of Lashley's diverse cerebral ablation studies, using rats. Sperry's work on rotating the eyes of male frogs and so producing striking reversals of space perception is also discussed, as are Penfield's investigations into precise parts of the human brain undergoing electrical stimulation. Aside from the physiological work, Zangwill illuminates the experiments of Michotte on the perception of causality, and those of Bartlett on skills.

The following article, 'Aspects of Psychopharmacology', by Hannah Steinberg, presents a short review of studies which have explored effects of diverse drugs on emotion, efficiency of test performance, and physiological activity.

Two chapters on perception follow: first, 'Attention and Visual Perception' by M. D. Vernon, then 'Perceptual Organization and Action' by James Drever. The former deals with role of attention in arousing and directing perception to some particular aspect of the environment. The subcortical reticular network seems to provide the neuro-physiological basis here. In the article by Drever, cogent argument is made for more thorough investigation of motor activity background during any experimental study on perception.

Chapters 11 and 12 are concerned with thought processes. 'Fifty Years' Experiment on Thinking: 1910–1960' by Prof. Humphrey, discusses the roles subsumed by 'association' and 'motive' as guides for and initiators of thought processes. 'Thought and Machine Processes', by Dr Bowden, compares the properties of human thought processes with those of electronic computers. Reception of redundant information is used by the brain to ensure an accurate picture of external events; machines can make little (if any) use of redundancies.

'Information Theory in the Study of Man' (by D. M. MacKay) is the first of three chapters on what may be rather broadly considered as 'problems of communication'. The structure, and flow, of information is clearly presented; enough to make one doubt, MacKay's italicized remark (p. 221) that 'Any reduction of redundancy brings a corresponding risk of failure to communicate'. Focusing more specifically on visual and auditory discrimination, the article by Prof. Summerfield and David Legge, 'Information Theory and Perception', shows how information theory may be applied to determine (in various experiments) the number of 'bits per stimulus' and/or 'bits per second' that a subject can transmit. In his chapter, 'Contact Between Minds', John Cohen discusses an approach which deals with problems of communication on a social-psychological level; Cohen compares scientific groups with other (e.g. military) groups. Science groups function to comment on, criticize, and publish the research work and ideas of scientifist; such groups thus '...may become vehicles of innovation'.

Next (in chapter 16) I. M. L. Hunter writes on 'The Organization of Memory'. Most of what we remember is based on schematization of meaningful items rather than storage of meaningless material. Following Hunter's chapter, P. E. Vernon contributes a characteristically lucid article on 'The Psychology of Intelligence and G'. The author contrasts operational, biological and psychological theories of intelligence.

Chapter 18 is entitled 'The Sources of Emotion'. Paul Fraisse (Professor of Experimental Psychology at the Sorbonne) discusses the neuro-vegetative reactions of a subject who cannot respond to the situation in an adaptive manner. The effects of task difficulty, motivation and subject's personality are dealt with in some detail.

Lack of journal space permits no more than mention of the last five chapters; as the editor says in his preface, these chapters '...are concerned with rather more specific themes:...'. 'Response to Stress in Military and Other Situations' is by D. E. Broadbent. 'Temporal Integration and Behaviour' is by Prof. Hearnshaw. 'Research Techniques in Psycho-Analysis and in General Psychology; an essay in Contrasts' is by Edward Glover. 'The Effective Use of Manpower' is by Prof. Rodger, and the final chapter 23, 'Some Concepts of Ethology', is by W. H. Thorpe.

H. WALLERSTEIN

Experiments in Psychology. By S. Jalota. London: Asia Publishing House. 1964. Pp. xii + 157. *27s. 6d.*

Practical classes provide the focal points of both general and honours courses in psychology. There are, broadly, two ways of designing such courses. The first of these derives from the European tradition of University teaching in which, it is claimed, the onus of thinking is placed squarely upon the undergraduate. The second has developed within the framework of American education in which courses and syllabuses are carefully planned, deviations and departures from them being minimal. The European method remains in use in some British Universities. Here, students working individually, or in small groups, largely define their own problems and are encouraged to devise a variety of solutions to them. While the laboratories in which such classes are conducted might inspire Emmett, they nevertheless become difficult to equip and supervise, frequently taking on an apparently chaotic and forbidding aspect. The American method, in contrast, is a much more tidy procedure. Here students are provided with pro-formae neatly listing hypotheses to test, comprehensive references and statistical techniques along with partly or wholly made-up sets of apparatus, diagrams and instructions as to procedure. Each of these two methods has features to commend it. The American method ensures that students make a systematic coverage of the basic fields of experimental psychology and learn to apply experimental techniques to probably as many as fifty or sixty problems. It has the disadvantages of frequently appearing contrived and repetitious but these may be overcome in a well designed course. The European method allows students to take a variety of problems to some depth and to be introduced to 'real life' questions at a relatively early stage. It has the disadvantages of being difficult to ensure that a comprehensive range of techniques has been encountered and that a relatively smaller number of studies will be made during the career of an undergraduate, but again these may be overcome with skilful teaching.

While leaning appreciably toward the American model, Prof. Jalota has, in this volume, made some attempt to steer a course between these two approaches. He lists forty-nine experiments, dividing these almost equally between 'Qualitative' and 'Quantitative' studies. In each case the format is standard, the problem stated, the required materials listed and the procedure described. The scope of the experiments ranges from the 'classics' of determining colour zones, locating the blind spot, measuring reaction times and mapping touch and cold spots; through measurement of illusions, consideration of parapsychological phenomena and threshold estimation; to the application of experimental procedures in constructing test instruments and a consideration of the structures of social groups.

University teachers responsible for arranging practical courses may find that a number of these experiments lend themselves to inclusion in their repertoire and, with the exception of the statistical procedures, which are inadequately described, first-year students will find little difficulty in carrying them out. Many readers will, however, be disappointed in a number of features of the book. While the majority of experiments are variations of well-known investigations, the only references given are to early editions of seven general texts, three at least of which are out of print. Many problems are stated and experiments conducted in such a way that we frequently appear to be involved in providing trite answers to naive questions. Never do we consider hypotheses, infrequently do we realize that we may pass on to related problems and often we are taken aback by such statements as '...*Factor analysis may be attempted, if desirable*'! (p. 142). Further, the distinction between 'Qualitative' and 'Quantitative' studies quite escaped the reviewer.

In spite of these considerations, Prof. Jalota has made a serious attempt at a task which is probably beyond the scope of a single person. The range of problems, techniques, methods of design and analysis which requires consideration in a modern manual of experimental psychology is almost certainly a task for a consortium of authors.

GERALD H. FISHER

Aspects of Applied Psychology and Crime. By A. A. ROBACK. Cambridge, Mass.: Sci.-Art Publishers. 1964. Pp. 326. $7.50.

This book is described as an unusual presentation and discussion of hundreds of little-known facts in human behaviour and in human misbehaviour, with practical recommendations and a minimum of theory.

Dr Roback has been one of the most influencial writers on problems of personality and character in the present century. Several of his books have been through a number of editions and translated into a variety of languages. His well-known *Psychology of Character* is recognized as a classic and has been brought out in a new and enlarged edition recently. This is partly because, unlike many present-day writers, he can express his thoughts in clear and readable language, and not after the fashion of a 500-page telegram in no particular language. Much more, however, it is because of his outstanding ability to extract essential facts, concepts, interpretations and evaluations from many sources, and to present them to the reader in a comprehensible scheme. The reviewer might say that he owes any understanding he has of personality and character entirely to his reading of Dr Roback's books, and this does not mean that it is limited by them, because Dr Roback always succeeds in opening the way to wider fields and gives a perspective which none can gain simply from lists of test-data. In other words, Dr Roback's approach is a human-oriented rather than a test-oriented one, and this is why his books have been so widely appreciated.

The present work is a collection of thoughtful essays and papers, many of which have appeared in journals, such as the *Journal of Applied Psychology* and the *Journal of Philosophy, Psychology and Scientific Method*. Two of the early chapters were actually papers read to Münsterberg's seminars when Dr Roback studied under him at Harvard, and still have life.

Many persons, academic and otherwise, concerned with philosophy, psychology and education, will read this book with great interest, even if not always to agree with it. R. W. PICKFORD

Clan, Caste and Club. By FRANCIS L. HSU. London: Van Nostrand. 1963. Pp. xi+335. 62*s*.

Dr Hsu, who is head of the Department of Anthropology at Northwestern University, has for a number of years been engaged in the endeavour of giving an account, an explanation even, of the patterns of culture of large societies in terms of their kinship organizations. This book is an extension of his *Americans and Chinese: Two Ways of Life*, resulting from two years work in India. It is a broad and exciting analysis of Hindu India using the comparative method, with Pre-Communist China and the United States as the other points of the triangle.

Hsu's primary data concern the family. He then moves to wider social groupings that are formed to give individuals the *élan vital* of existence—sociability, security and status. The correlation of kinship organization with ways of life is not a new essay; the culture-personality school for one, were here before him. Hsu suggests that their work lacked finesse, and failed to convince anthropologists because their conceptualization has been only in terms of kinship *structure*. He has introduced the variable of kinship *content*. Whereas structure describes the features that govern the formal arrangements between individuals standing in reciprocal categories, content refers to the characteristics that govern the tenacity, intensity, and quality of interaction between people related by kinship. Structure and content are not entirely disconnected. 'The content of a kinship system is to a great extent determined by the emphasis given one or another particular primary relationship in the kinship structure.'

The observations and argument of the book may perhaps be summarized in the following way. The Chinese social organization is dominated by the father–son relationship which leads to a situation-centred orientation of culture, expressed in the form of mutual dependence. For the Chinese his family and its extension the clan were the beginning and end of his human universe. He was convinced of the permanence of his place in the clan. The centripetal mode thus fostered produced kinship solidarity, standards of mutual obligation, and social stability of outstanding strength. The Hindu social organization is dominated by the attributes of the mother–son relationship, expressed in the form of unilateral dependence. The temporary quality inherent in an attachment to the mother leads to a wider centrifugal mode. The only higher aspiration and point of reference lies in the supernatural sphere. A seeming boundless unilateral dependence is expressed in the Hindu religion, and caste is the only social refuge. The husband–wife unit is pre-eminent in American society, which feeds an individual-centred orientation, and in turn leads to the high value placed on self-reliance. Consonant with the latter we find social groups based on individual choice and some form of contract, that is clubs.

In dealing with large social complexes such as these, the data must obviously be of a different order to those of ethnographic studies. Like Ruth Benedict in her account of Japan, Hsu deals here mainly with what Redfield has called Great Traditions rather than Little Traditions—with history, tradition, ideological and theological precepts, and not only with particular examples of the functioning of modes. His discussion of the differences between caste in India and caste in the United States, for instance, demonstrates the value of this scheme. The danger attending

Great Tradition thinking is clearly that of over-simplification; this has not entirely been avoided. Clan was undoubtedly the principal social organization in the Great Tradition of China and functioned for the gentry; however, Hsiao-tung Fei has made much of the point that it was not available in the restricted life of the peasantry, social organization and obligation stopping at the expanded family and the neighbourhood.

An examination of further expressions of clan, caste, and club is planned for a future volume. Details of religion, art, philosophy, technology and politics will come there. However, even in the present over-view it would appear that the advent of Communism in China merited greater space. Such platitudes as that the success of the new ideology depends on whether the government is capable of learning from its mistakes are unworthy of the book.

HALLA BELOFF

Expression of the Emotions in Man. Edited by PETER H. KNAPP, M.D. New York: International University Press. 1963. Pp. 351. £3. 8*s*.

This is another multi-authored American book on the emotions, originating from a symposium held in New York in 1960. Every symposium seems, nowadays, to spawn a book. The names of the same contributors turn up with some regularity, saying very much the same thing in different surroundings. As one of the participants complains: 'Dr Spitz's presentation contains so many different ideas and touches upon so many empirical issues that it will be impossible to do the paper justice.' The present reviewer echoes these sentiments as applied to the whole book.

The well-known authors present potted versions of their original contributions, most of which have appeared in potted versions elsewhere. Discussions are also included of the papers, and these do contribute something useful, but nowadays most of these 'discussions' are printed from abstracts prepared before the Conference.

However, these difficulties apply to all endeavours to build up communication between disciplines, something that very much needs doing if each branch of science, in an age of ever-increasing specialization, is not be become entirely unintelligible to the others. Certainly the level of individual communications from authorities such as Gregory Bateson, Margaret Mead, David Hamburg, Karl Pribram and Rene Spitz is high enough, within the limitations prescribed by this medium. Every aspect of emotions—physiology, psychology, psychoanalytic, linguistic, social and anthropological—is touched upon.

J. R. SMYTHIES

Cognition: Theory, Research, Promise. Papers read at the Martin Scheerer Memorial Meetings on Cognitive Psychology. Edited by CONSTANCE SCHEERER. London and New York: Harper and Row. 1964. Pp. xi + 226. 30*s*.

This title reminds one inevitably of the problem which John B. Watson tried to persuade himself he had solved fifty years ago. And he had indeed persuaded many young workers, who continued at the same time to talk of 'the colour yellow' and the protanope.

For in spite of 'The Behaviorists' and the ingenious verbal gimmicks they contrived, Cognition, Awareness, 'experience subjectively regarded' remain with us, intractable but indispensable, and this smallish book gives a sample of a stratum of ingenious work done on them. The book is by no means easy to read, but the reviewer found it exciting. 'No one' says Taft, the *rapporteur*, 'has offered us a definition' of cognition. Strangely, this hardly seems to matter. For a long time nobody offered us a definition of electricity. There are eight papers, with a short introduction on Martin Scheerer by Heider and a Summation by Ronald Taft, of the University of Western Australia. The reviewer took to one by Witkin, on 'The Origins of Cognitive Style', which he found pleasantly imaginative, and which, for example, correlates certain traits (may one call them so?) of perception with intellectual and family function and maternal relationship. Rock and Ceraso's chapter on 'A Cognitive Theory of Associative Learning' is critical and perhaps destructive of certain parts of conventional associative theory, and is founded on rather disturbing experimental results—disturbing, that is to the conventionally minded. But without detailed analysis it is invidious to single out individual chapters. The book is recommended to psychologists who do not mind their faith being mildly upset, and is full of suggestions for research.

GEORGE HUMPHREY

Symbol Formation. By H. Werner and B. Kaplan. London: John Wiley and Sons. 1963. Pp. xiii + 530. 90*s*.

A book review is the converse of a piece of consumer research, for instead of having a variety or products and a fairly standard consumer one has a completely standard product and a diversity of potential consumers. It is unfortunate that this makes the reviewer the sole source of error in the appraisal, for particularly with a book of this sort reactions are likely to vary from critical interest to impatient rejection. The central ideas in this book are concerned with the relations between symbolization, inner and outer speech, thought, imagery, dreams, gesture, psychopathology, and non-verbal representation. The conceptual framework is descriptive rather than explanatory, and one feels that the evidence could be equally well presented to support a different theoretical approach. The authors have their own peculiar style of presentation, and the reader's task is made difficult because of an excess of technical terminology and a lack of regular summaries. The best parts of the book are probably those on language in schizophrenia and on the analysis of internal and external speech in sections III and IV. The book may appeal to a small number of psychologists who are deeply involved in the study of language and communication, but it is not a book for the general reader and not the best introduction to the psychology of language. But it is a thoroughgoing statement of a special point of view about the formation of symbols, and the authors are so knowledgeable about the subject-matter that the interested expert cannot afford to ignore their contribution. There is rather more argument than information in the book, and one gains the impression, at least from the earlier sections, that the experiments are weird and unsatisfying; they leave too much room for interpretation, so that the evidence and examples often appear to be selected to illustrate and support the theoretical argument. The bibliography contains 284 titles, with very useful references back to citations in the text.

D. B. BROMLEY

The Physiological Basis of Mental Activity. Proceedings of A Symposium Held in Mexico City, 1961. Edited by Raúl Hernández Péon. London: Elsevier. 1963. Pp. vi + 283. 90*s*.

This is a well-produced volume of nineteen specialized papers by distinguished research workers. The papers are all of a high standard and report original research into the physiological basis of central sensory mechanisms, sleep, learning, motivation, and social behaviour, although half the papers are concerned with sleep.

Mancia finds spatial and temporal differentiation of evoked potentials recorded from electrodes in the olfactory bulb, for at least two categories of odours. The evidence for centrifugal control of activity in the bulb is discussed.

Brooks, in a somewhat loosely constructed paper, gives a review of work in his own laboratory and the work of others on afferent convergence and divergence in relation to variability and redundancy in cortical function.

Reinoso-Suárez finds that lesions in the subthalamus and mesencephalic reticular formation cause a decrease in retinal potentials evoked by light and in the *b* wave of the electroretinogram. But whether this control is via centrifugal fibres in the optic nerve or via a sympathetic vascular pathway is not known.

In several papers, the electrical responses to a series of similar stimuli are recorded from single units or from surface electrodes and automatically averaged and analysed so as to bring out small responses which are time-locked to the stimulus and which are otherwise hidden in neuroelectric noise. Rosner, Goff and Allison use this technique to analyse a complex of responses in the cortical hand areas of man and monkey to stimuli applied to the hand. Analogous surface-positive waves are found in the two animals these corresponding to the activity of the reticular system. In a second paper the same group extend these findings by the use of an anaesthetic drug. Brazier proposes that the profile of averaged activity in an aggregate of neurones is an important information-coding device in the nervous system. This is, of course, not a new suggestion. but Brazier makes a modest start towards giving it an experimental basis.

Garciá-Austt reports that the cortical potentials evoked by photic stimuli in man and rat are affected by changes in attention. Negashi, Bravo and Verzeano use convulsant agents to study the relation between neuronal activity and gross activity.

It has become evident in recent years that the reticular system has a depressive as well as an arousal function. Monnier, Hösli and Krupp confirm that low-frequency stimulation of the reticular formation depresses cortical activity and behaviour, but that high-frequency stimulation causes a general arousal. Rossi also confirms these findings but finds that low-frequency stimulation produces light sleep, not merely depression of activity. He argues that the centre for light sleep or so-called telencephalic sleep is in the caudal pons and/or medulla. A centre for deep or rhombencephalic sleep is put between mid- and rostropontine levels along with cells having arousal properties. In contrast to these views, the importance of cortical centres in light sleep is stressed by Clemente and Sterman. It is suggested that a forebrain system acts via the diffuse thalamic projection. Hernández Peón and Chavez Ibarra report similar findings. One wonders to what extent these different suppressor centres are parts of one and the same system.

Jouvet and Jouvet find that dreaming in man is associated with light sleep, electroretinogram activity, and eye movements, and that similar sleep patterns occur in the cat. Buendia, Sierra and Segundo report that cats may be conditioned to wake selectively to a particular tone, even from deep sleep. Creutzfeldt and Meisch report on the relation between blood sugar level and electroretinogram activity in *encéphale isolé* cats.

Both Miller and Olds give reviews of studies from their laboratories on reinforcement mechanisms. Scheibel and Scheibel report on the stages of development of the cat's electroretinogram rhythms and relate these stages to anatomical changes.

Finally, Delgado gives an account of the effects of brain stimulation on the social behaviour of monkeys. Stimulation of the motor cortex and cerebellum produces effects which have no direct relevance for social behaviour and one wonders what Delgado expected in these cases.

Altogether, this is rewarding reading for the specialist and graduate student. It is reassuring to find that in several cases different investigators report essentially similar findings.

IAN P. HOWARD

•

Perception and the Conditioned Reflex. By YE. N. SOKOLOV. London: Pergamon. 1963. Pp. x+309. 80*s*.

English readers will have already seen an outline of Sokolov's work in Luria's London Lectures published as *The Role of Speech in the Regulation of Normal and Abnormal Behaviour* (Review in this *Journal* (1962) **53**, 90). This particular publication reports extensive work on the physiology of the orienting reflex. Pavlov originally observed that orienting or 'What is it?' reflexes interrupted the development of conditioning, and various workers have since drawn attention to the role of non-specific signals in perception. Sokolov, in his work with Olga Vinogradova, has shown not only that blood flow in the fingers and the external carotid artery vary inversely and concomitantly in orienting and defensive situations respectively, but has gone well beyond this relatively simple statement. He shows that the alerting of the organism which is revealed by orienting phenomena is accomplished through the control of blood flow which is connected, in his view, with excitation. Both orienting reactions to light and darkness (figure 36, p. 111) and the basic phenomena of the relationship between defensive and adaptive reflexes (figure 17, p. 67) are set out in terms of plethysmographic, eye-blink, E.E.G. and P.G.R. measures. The interactions of these measures are interpreted in various ways to suggest, for example, that the state of the organism rather than the intensity of the stimuli is the relevant variable in determining the effectiveness of such a stimulus.

In determining and controlling sensitivity in this way, orienting, adaptive and thermoregulatory reflexes interact. The technique generally used by Sokolov to study these interactions is a polygraphic one, supplemented by photography of the eye, for example. Sokolov claims that depression of the alpha rhythm is closely associated with the P.G.R. as a measure of orientation, but states in addition that the alpha can be suppressed by the onset of darkness in a well-lit room. This fact he uses to justify his view that alpha changes are part of the orienting response. A second point of interest is the effect of verbal instructions or the re-emergence or disinhibition of orienting reflexes. Verbal instructions to respond to a stimulus may produce orienting reflexes in the place of most established defensive ones.

Scientific Russian translates poorly and the book is so full of detail that the reasonably clear framework outlined in the introduction becomes unrecognizable. As a result one can scarcely begin to criticize methodology because little effort is made to present it. A further problem is that no concessions are made to the western reader who usually does not know his Pavlov very well. In view of these innate difficulties I think that the editors and the translator have done a useful service in publishing this interesting material.

N. O'CONNOR

Complex Human Behaviour. A Systematic Extension of Learning Principles. By ARTHUR W. STAATS and CAROLYN K. STAATS. New York: Holt, Rinehart and Winston. 1963. Pp. xii + 546, $7.50.

It is perhaps only with our comparatively recent concern about the development and uses of language, and of symbolic processes generally, that psychologists have given notice that this subject is coming of age. The painful choice between being either rigorous and scientifically irreproachable about trivial problems or being scientifically suspect in our approach to important ones is now much less of a dilemma for the practising psychologist. The work of Mowrer, Osgood, Skinner and Kelly provides us with conceptual tools with which human behaviour, perhaps not yet at its most sophisticated, but certainly at a complex and undeniably human level, can be investigated.

It is the achievement of this admirable text that it presents for the not too advanced student a readable introduction to the analysis of semantic behaviour. The authors begin with a short discussion of scientific method and follow it with a sound exposition of learning principles, which, although it relies heavily on Skinner's work, never becomes narrowly sectarian. On this foundation the authors present the real substance of their book in the two chapters dealing with language development and language function.

The six chapters that follow present us with glimpses of how useful this approach can be when applied to a refreshingly wide variety of human problems. In particular, there are discussions of personality, motivation, social interaction, child development, programmed learning, and a final section on behaviour therapy and the principles underlying it. In this second half of the book the professional psychologist may find little to excite him in the section dealing with the field in which he has special knowledge, but it would be surprising if he found nothing to stimulate his interest in the other sections. However, it will be the second- or third-year psychology student, the student of education, or the teacher in training who gains most from this book; for them, and for those of us who require occasionally to be reminded of the importance of verbal behaviour, reading this book will be a very worthwhile experience.

J. DREWERY

Children Growing Up. By JOHN GABRIEL. London: University of London Press. 1964. Pp. 368. 30*s*.

This book, intended for practising teachers and students in teachers' training colleges, is based on wide reading, as well as on experience in teaching children and the future teachers of children. Continuity of contacts with children and young people of both sexes clearly enabled the author to appreciate the needs of both, and to avoid the temptation of over-attachment to any one psychological theory of personality development. He does not hesitate to set out his material in text-book form, numbering and italicizing the points he particularly wants his readers to remember ('Dispositions may be classified as follows', etc.). All that he has to say is sound and uncontroversial, for he presents most of the findings of depth psychologists as they perhaps ought to be presented to students—as hypotheses, not as established truths. He quotes as freely from the older writers, such as McDougall, as he does from contemporary American research on character and personality. Concepts taken from the writings of Freud, Jung, Suttie and Susan Isaacs are explained in everyday language, calculated not to startle but to elucidate and inform. His exposition of the stages of infant development, however, is based on the studies of Gesell and Piaget.

Practising psychologists will not find anything but familiar material in this book, but they will welcome so much useful information welded into a whole, simply and clearly presented. Eminently sensible suggestions are made to parents concerning the wisdom of tolerating the inevitable differences between themselves and their adolescent children, and to teachers with regard to not attempting to be psychologists, psychotherapists and teachers, all in one. It is a wise book in that it lays emphasis on exceptions to the rule, on the complexity of every personality, and on the danger of labelling according to text-book classifications. Its aim is to enlarge the students' mental and emotional horizon and to increase their insight and sympathy, without encouraging an illusion of omniscience, an aim in which I think it succeeds.

LYDIA JACKSON

Children Tell Stories—an Analysis of Fantasy. By EVELYN G. PITCHER and ERNST PRELINGER. New York: International Universities Press. London: Bailey Bros. and Swinfen. 1963. Pp. 256. 36*s*.

This book is the product of a simple idea treated with academic distinction. During the years 1955–58, 360 stories were gathered from children aged 2–5 years (70 girls, 67 boys), mostly during regular nursery school and kindergarten hours, in New Haven, Connecticut. The first chapter has useful things to say about the functions of fantasy in general adaptation to life. The second describes the experience of obtaining the stories (clearly with skill and patience) after an attempt with picture-story methods had produced only naming responses from the youngest children. It is specially interesting to have stories from 2- and 3-year olds,.who are so seldom interrogated about their thoughts, and also to find that the 5-year olds were true to the picture of social conformity given to us for this age by Gesell; they were eager to co-operate but the authors had to overcome their tendency to report a story which they had heard or seen on television.

The stories form the long central chapter of the book and are happily presented without accompanying text, apart from being grouped according to age and sex.

The stories are analysed for formal characteristics on seven dimensions (e.g. expansion, differentiation of the main figure, complexity of characters, range of activities), with a summary of age trends. This kind of approach has been used elsewhere in a developmental study of 'Worlds', and in an interesting foreword to this book Dr Lowenfeld comments on the similarities between the stories and the contents of constructions made by older children with the World material.

The contents of the stories are next studied for developmental and sex trends in selected topics: aggression, parent relations, morality and other social themes. Erickson's (1950) formulation of developmental tasks (Trust, Autonomy, Initiative, Industry) is considered for possible validation from the trends of the stories. There follows an interesting chapter on Dynamic Themes and Defences, 'some of the less immediately manifest aspects of the stories' since the foregoing analyses have left the authors with a feeling of not having done justice to the richness of individual stories.

Workers in applied fields in this country may find the study too much oriented to theory (this is not of course a criticism of the book). The study is not specifically directed to the clinical use of the stories, and the authors set themselves virtuously against 'wild analysis' without closer knowledge of the individual children. On the other hand, there is no reason why those with the requisite skills should not use the ideas of this book in their clinical work. The recorded stories themselves will give pleasure to academic and non-academic alike, and will surely arouse eagerness to collect more stories.

L. R. BOWYER

The Perceptanalytic Scale. By ZYGMUNT A. PIOTROWSKI and MILTON R. ROCK. London: Grune and Stratton. 1963. Pp. iv + 220. No price given.

This book describes an attempt to study the potentialities of the Rorschach ink blot test for the identification and selection of top executives for industry.

The study is based on a comparison of successful and unsuccessful top executives, their success or failure being judged by their actual performance in their current positions. Only those executives were studied who occupied the highest rank of managerial responsibility. Special methods of scoring the test were developed for the purpose of the study, those test items being selected which seemed most closely related to the capacities of executives.

The authors were satisfied that there were significant differences between the test scores of the successful executives and those of the failed executives when the tests were scored by their own specific method. Their presentation of their results is also convincing to the reader.

However, the authors are rightly modest about the implications of their study. They are aware that their sample of executives may have been biased and that therefore one may not be able to accord universal significance to their results. They are anxious that further validatory studies should be carried out. Further, the success of the application of this method depends on both the skill with which the test is applied and on the evaluation of the results in a way consistent with that used by the authors, criteria which may be difficult to achieve on any wide scale.

Still another problem presents itself, which this book does not really tackle: that is, the matching of a specific potential top executive to a specific top executive position. In other words, even if it is possible to show by using the perceptanalytic scale that a candidate has top executive potential, this study does not deal with the question of whether he could realize his potential in any specific position. The study makes clear that successful top executives may be very different as people. The organizations in which they may hold top executive positions are as different. This book does not tackle the even more complex task of defining the essential features of an organization into which the top executive must fit. The crucial task for selectors is the matching of the idiosyncratic good performance of the man with the idiosyncratic features of the organization in which he may be placed.

One may perhaps hope that the authors will give more attention in future work to this aspect of the problem.

ISOBEL E. P. MENZIES

Concepts of Personality. Edited by JOSEPH M. WEPMAN and RALPH W. HEINE. London: Methuen. 1964. Pp. xxix + 514. 70*s*.

This is a collection of seventeen essays by different authors, partly based on a series of lectures in the University of Chicago. The essays vary considerably in apparent aim, in clarity, in level of sophistication and in originality, as is probably inevitable when the common theme is so broad and the area of psychology so patchily developed. The average standard is high, and although few readers will be entirely satisfied with the book as a whole, most will find much new and interesting material.

To the present reviewer, the most rewarding sections, in the order in which they appear, are: (1) Eriksen on 'Perception and Personality', containing a competent survey of work on perceptual vigilance and perceptual defence. (2) Butler and Rice on 'Adience, Self-Actualization and Drive Theory', bringing together the ideas of Maslow and Rogers on personality development, the results of experiments on sensory deprivation, and work such as that of Berlyne on exploratory behaviour. (3) A characteristically difficult contribution by Cattell on 'Concepts of personality growing from multivariate experiment', full of 'exmodulators' and 'path experience coefficients' and of important and original ideas. (4) A useful article by Fiske, largely concerned with the special problems of validity that arise in personality assessment.

To what kind of reader is this book most useful? On the whole it can be recommended to the student who is already familiar with, say, Hall and Lindzey's 'Theories of Personality' and who requires a second, still general, text that is less systematic and more advanced. He will find much in this collection that is suggestive and stimulating.

H. J. BUTCHER

Experiments in Motivation. Edited By H. J. EYSENCK. Oxford and London: Pergamon. 1964. Pp. vii + 424. £5.

This book contains a series of review and experimental report articles all of which are concerned with the topic of human or animal motivation. The book is divided into two parts, Part I being concerned with human motivation and Part II with experiments with animals. Part I is further divided into two sections, Section A consisting of reports of a series of experiments designed to explore differences in performance between a high drive and a low drive group, the drive level being fairly convincingly determined by situational manipulation, and Section B consisting of three articles concerned with ego-involvement and attitudes as they affect motivation.

Whatever strength the book has lies in the relatively wide range of clearly presented experimental evidence, particularly in Section A of Part I and in Part II. Its main weakness is the rather inadequate and sometimes positively naïve theoretical appraisal of the field, which is based on a somewhat barren neo-Hullian system or upon an imperfectly understood physiological foundation.

The lack of any attempt to provide a coherent unifying theory might be regarded by some as an asset, but such potential readers are likely to be even more irritated by the piecemeal theoretical interpretations which are made than they would be by a general theory. Those who look for unifying concepts are equally likely to be irritated and disappointed, particularly as no real effort is made to relate the reported findings to the wide range of very relevant neurophysiological evidence on drive which is currently becoming available.

Presented simply as a set of research reports this book could have been much shorter, possibly much cheaper, and probably more useful. As it is, it gives the impression of being hurriedly conceived and put together. Several articles, such as Feldman's on motivation and task performance, could have been left out altogether as they add nothing to previous reviews of the subject, such as that by Duffy in her recent book *Activation and Behaviour* (review in this *Journal* (1963), **54**, 180–1), and serve in no way to clarify the reported experiments. For the same reason several other articles, notably those in Section B of Part I, should have been much more stringently edited, as they contain remarkably little new useful information.

In spite of the fact that it does contain a certain amount of interesting and new experimental evidence, the amount of useful content seems to offer little justification for a book of this size.

R. S. CORTEEN

Social Behaviour and Organization among Vertebrates. Edited by W. ETKIN. Chicago and London: University of Chicago Press. Pp. xii + 308. £2. 16*s.*

In the editor's opinion 'an understanding of human nature must be founded upon an understanding of social behaviour in lower animals'. This book is a distinct contribution to such understanding. His five chapters constitute a book within a book and both lay the foundation for, and draw together, the specialist contributions of D. E. Davis, F. A. Beach, D. S. Lehrman, N. Tinbergen and J. P. Scott.

A major concern of this book is with the endocrine correlates of behaviour, especially that most social of behaviours, sexual behaviour, to which, together with reproduction generally, over a quarter of the whole book is specifically devoted. The editor provides in a single chapter a compressed course in endocrinology—so compressed, indeed, that the student without some prior knowledge might derive little benefit from it. Davis follows with a useful account of the psychoendocrinology of aggression, with special attention to crowding in animal populations. The editor's review of reproductive behaviour in birds and mammals is supplemented by a masterly treatment of its hormonal basis by Beach and an equally authoritative treatment by Lehrman of the cyclic control of reproductive behaviour.

The editor resumes his expository role to outline ethological theory in the next chapter, which is followed by a fascinating discussion by Tinbergen of the evolution of signalling devices (previously called 'releasers') and the selection pressures which are believed to give rise to them. Scott's chapter on the effects of early experience on social behaviour reviews the infra-human evidence and provides a tentative application of it to the development of the human infant. The editor's final chapter contrasts the types of social organization found in birds and mammals and returns to the theme of the evolution of behaviour, this time in connexion with man. Here Etkin sustains a high level of discussion which is at once stimulating and satisfying.

Perhaps it is inevitable, in view of the numbers of contributors, that one's overall impression of this book may remain somewhat diffuse, despite the editor's frequent cross-referencing between chapters. But the uniform excellence of the many illustrations which adorn it throughout, a large proportion of them most sensitive renderings of behavioural sequences derived from photographs, more than compensates and gives further cause to welcome it as a valuable addition to the growing number of up-to-date texts one can recommend to the student of comparative psychology.

P. L. BROADHURST

Maternal Behaviour in Mammals. Edited by HARRIET L. RHEINGOLD. New York: and London: Wiley. 1963. Pp. viii + 349. 66*s*.

The mammals reported on in these original papers are laboratory rats, deermice (and other species of *Peromyscus*), rabbits, cats, dogs, sheep and goats, moose and elk, rhesus monkeys, langurs and free-ranging baboons. The authors are mainly, but not all, psychologists, and all currently researching in America, in the laboratory or the field, on the animals which they write about. The choice of animals is obviously determined by which are being studied at the moment, and Prof. Rheingold is fortunate to have collected such an interesting variety.

Psychologists and others working with human infants often draw analogies with infra-human maternal behaviour, and until now they have lacked a source book. (The relevant sections in E. S. E. Hafez: *The Behaviour of Domestic Animals*, London, 1962, provided the nearest thing to such a compendium.) From this point of view the book is invaluable. One of the things which it makes abundantly clear is that the comparative method must be applied with caution. Prof. Rheingold points out that, even within a single order, such as rodents, there is great variability between species in the degree of maturity of the individual at birth. 'How much more caution should be exercised before generalizing from genus to genus and from order to order!' But despite such variability there are common themes which would be found also even in a different class of vertebrates, the birds. (Here the Hafez book has the edge, since it includes some birds.)

Apart from comparative purposes, the chapters of this book are individually of great interest. Particularly those on the rat (Rosenblatt and Lehrman), cat (Schneirla, Rosenblatt and Tobach) and dog (Rheingold) are comprehensive and contain details of elegant experiments.

If a prediction is to be made from this book, it might be that research will focus more on the period when maternal behaviour ceases, through the onset of punishment by the mother, and by her withdrawal. As Harlow puts it in his chapter on the rhesus: 'From here on mother does not care for baby fingers in her hair; a touch that once went to her heart now merely makes the hair depart.'

B. M. FOSS

Perceptual and Visuo-motor Disorders in Cerebral Palsy. By M. L. J. ABERCROMBIE. London: The Spastics Society and Heinemann. 1964. Pp. 136. 25*s*.

This study was commissioned by the Medical Advisory Committee of The Spastics Society on a topic of current lively interest. If is the outcome of five years' survey of the literature and personal engagement in research by Mrs Abercrombie. The subject has been much in need of clarification from its theoretical, clinical and educational aspects. The present monograph therefore renders a valuable service to all those interested, whatever the length of their acquaintance with the phenomena discussed.

The treatment falls neatly into two halves. The first deals with the nature and incidence of the disorders, their relation to other handicaps, the strategy and tactics of research, and tests. The second consists of abstracts of some of the literature, referred to in the previous text.

The disorders belong to that group whose occurrence after neurological damage in adults is not seriously disputed, although much of the evidence comes from individual case-studies, but whose diagnosis in educationally backward children is regarded with some suspicion by the statistically oriented. This study will not resolve these doubts, but it does bring out clearly the grounds for differences of opinion, and appends the available evidence.

Mrs Abercrombie adopts explicitly the standpoint that the disorders are 'specific'. She considers that this is best demonstrated by regarding the child's performance in relation to its '*verbal* mental age'. She makes it, however, clear that this is clinicians' specificity and not statisticians', quoting papers which failed to demonstrate an important group factor of this type.

It might be remarked that there is a strong tendency, when the psychological function of brain-injured children is considered, to adopt verbal facility as the base-line from which deviations are measured. It is certainly usually (pace Hebb) their best level. With brain-injured adults, on the other hand, merely the various discrepancies are noted, and verbal level would not be used to estimate capacity for learning new skills. The former practice entails explanation of subsequent poor educational progress in terms of specific disabilities rather than a general one.

In this field there are more problems than solutions, and much remains to be investigated of practical as well as theoretical importance. As Mrs Abercrombie points out, more longitudinal studies are required to establish the relationship between educational achievement and 'visuomotor' test performance. She also stresses the distinction between visual perceptual and visuomotor disorders, the relative lack of information about the former in this field, and the variety of possible explanations of the latter in children already selected on the basis of disordered motor function.

She also stresses the possible importance of disorders of eye movement, in particular squinting, in the genesis of the condition, as suggested by the research in which she has been concerned. Unfortunately, few of the other studies quoted could be relied on to provide reliable information about this in their subjects or their controls.

In general, it is visual perception that is discussed, although touch and 'body scheme' are also mentioned, and contrasted with language function. It is worth remarking that the differences involve more than sensori-motor modality. It is never suggested that children with 'visual perceptual' or 'visuo-motor' disorders have difficulty in visual recognition of ordinary objects. It is accurate perception and reproduction of unfamiliar shapes in terms of their formal rather than associative characteristics that is faulted. The contrast is made with overlearned meaningful words and not with meaningless sounds. Acceptable standards of success in reproduction also differ radically between the modalities.

It seems possible that much of the difference in task difficulty, and hence the apparent specificity, could stem from difference in degree of novelty and the child's capacity to cope with this, except when there is some neurological abnormality of sensorimotor function, such as Mrs Abercrombie suggests.

The wide discrepancies in aptitudes will, however, continue to present serious practical problems to teachers, whatever the theoretical explanation. Some help with curricular rearrangement, and in deciding whether to wait for time to mend matters or to go in for vigorous training, would be very welcome. At present, there are no unequivocal grounds for decision either way. There are good grounds for careful examination of the children if we are ever to find out.

J. NAUGHTON

Monographien aus dem Gesamtgebiete der Neurologie und Psychiatrie (*Monographs in the General Field of Neurology and Psychiatry*). Edited by M. MUELLER, H. SPATZ and P. VOGEL. Berlin–Goettingen–Heidelberg: Springer.

Psychopathen: Daseinanalytische Untersuchungen zur Struktur und Verlaufsgestalt von Psychopathien (*Psychopaths: Existential-analytic Investigations on the Structure and Formal Course of Psychopathies*). Nr. 94. By H. HAEFNER, with an introduction by L. BINSWANGER. 1961. Pp. iv+230. DM. 48.

Die Experimentelle Psychose—ihre Psychopharmakologie, Phaenomenologie und Dynamik in Beziehung zur Person (*The Experimental Psychosis—its Psychopharmacology, Phenomenology and Dynamics in Relation to Personality*). Nr. 95. By H. LEUNER. 1962. Pp. x+275. DM. 69.60.

Der Fruehkindliche Autismus: eine Klinische und Phaenomenologische Untersuchung am Leitfaden der Sprache. (*The Autism of Early Childhood—a Clinical and Phenomenologic-Anthropological Investigation Perusing Language as a Guide*). Nr. 96. By G. BOSCH. 1962. Pp. vi+123. DM. 36.

The three monographs reviewed belong to a long series, now nearing the hundred mark, which occupies a unique place in the historical development of continental neurology and psychiatry, unique in that it represents a fundamental re-thinking of specific clinical problems.

Haefner in his work on psychopaths employs and develops the phenomenological approach of the Heidelberg school. He is not concerned in deciding between the relative importance of heredity and milieu in the problem of psychopathy, but to arrive at an understanding and formulation of it on the basis of existential analysis, thus following Binswanger, who provides a searching introduction to the book. Haefner does not see the end of the road for the difficult problem of psychopathy in the formulation of a typology like Schneider's, but he aims at a fundamental comprehension of the psychopath's way of 'living in the world'. His conclusions are based on three thoroughly discussed 'paradigmata' of psychopathic life development. His concepts of 'rigid consequentiality of failure' to come to terms with the surroundings, of illusionary, obstinate or forcible self-'fulfilment' which objectively are rather self-defections, of the 'facade' encrusting and impeding a growing self—these appear to the author to hold out more promise than the mere statement of 'maladjustment'. Whether the richness and multifariousness of life and clinical experience can all be pressed into a grand design of this kind, remains open to doubt.

Leuner's work on experimental psychosis is meritorious in that it develops the rather static phenomenology of Jaspers into a more dynamic or—as the author calls it—a 'conditional-genetic' one. More than 1000 prolonged sessions were carried out with seventy subjects mainly using L.S.D. and to a less extent mescaline and psilocybine. The detailed protocols of some of the sessions make interesting reading and in his interpretations the author tries to bridge the gap between traditional clinical psychopathology and the psychodynamic approach.

Bosch, in his monograph, deals with childhood autism in a new way, i.e. not by collating phenomena, but by examining these children's way of life according to the principles of phenomenological anthropology. On the basis of his own thirteen cases and the literature he concludes that the rift in development occurs in the necessary step from primitive symbiosis to effecting the structuralization of 'I, You and the World'. This becomes very apparent in defects in the development of certain dimensions of language. Through imitation and automaton-like activity some of these children impress as more developed than they actually are in the decisive field of communication.

STEPHEN KRAUSS

Systems of Psychotherapy. A Comparative Study. By DONALD H. FORD and HUGH B. URBAN. London and New York: Wiley. 1963. Pp. vii+712. 83*s*.

This is a worthwhile venture which might provide a basis for new developments in thinking about psychotherapy. The authors have systematically analysed ten different theories of psychotherapy and have managed to do so without being tendentious. Unavoidably, protagonists of Freud's psychoanalysis, ego-analysis, Dollard and Miller's learning theory, Wolpe's reciprocal inhibition, Adler's individual psychology, Rank's will therapy, Rogers's client centred-psychotherapy, existential analysis, Karen Horney's character analysis and Harry Stack Sullivan's interpersonal relations, will find errors of omission and emphasis in the accounts of their own models. The uncommitted reader, however, can get an informative and fair picture of all these approaches from Ford and Urban.

The book is in three sections. The first section is devoted to a long account of a system for comparing the ten varieties of psychotherapy and its methodological and theoretical implications. The second section is the core of the book, viz. the presentation in chapters of between 30 and 70 pages of the ten selected psychotherapies. In this section the initial ponderous concern with how to compare systematically pays off in a restrained and standard description of each method. The third and final section contains a discussion of the similarities and differences in the ten theories, which remains within the confines of the initial criteria for comparison, and will disappoint any reader who expects to have his thinking done for him and to be told the 'best buy' as in *Which?*. In spite of their objectivity the authors' preference for Harry Stack Sullivan glints through occasionally.

The pragmatist will be disappointed that effectiveness of each treatment is not stressed in the systematic comparison. This is a reflexion of the confusion that currently exists in psychotherapy and which will not be resolved until someone makes public the extent to which psychotherapy does good. Like others in the field, Ford and Urban are coy about the notion of effectiveness. 'To ask whether therapy is effective is a misleading question. A more useful question to ask is: what kinds of treatment procedure will produce what kinds of behavioural changes in what kinds of individuals?' But this is surely all that is meant by effectiveness, and until the material is available for this practical proof, comparing systems of psychotherapy will remain a slightly desiccated pursuit. Granting this defect, which the authors can only be criticized for condoning, the book is valuable as a balanced statement of the systems. Students of psychology can use this as a source of reference for the various brands of psychotherapy, and, with supervision, could devise their own bases for comparison. The psychiatrist in training could profitably consult the description section as a hand-book of psychotherapy. R. M. MOWBRAY

Freud: a Critical Re-evaluation of his Theories. By REUBEN FINE. London: Allen and Unwin. 1963. Pp. 271. 35*s*.

This book attempts in small space to carry out a highly important task: to work systematically and chronologically through the writings of Freud, and trace step by step the cautious and continuous development of his theories, and the changes of concepts that he actually and frequently made, stage by stage. It is impossible to discuss Freud without such knowledge, let alone to judge accurately of his work. So fertile was his mind and so constant was the impact of his clinical experience on his theoretical thinking, that he was constantly out-dating his earlier work, without always realizing the extent of this or admitting it explicitly. It is easy therefore to misrepresent Freud in subtle ways. Dr Fine cites Kardiner's report of a meeting of the Vienna Psycho-Analytical Society in which members were quoting Freud in his presence, till he protested: 'Here you are debating what I had said, and not one of you thinks of asking me "What do you think now?" If this is the way you treat me while I am still alive, I can well imagine what will happen when I am really dead.' Dr Fine's type of study is the only way in which we can now ask Freud 'What do you think now?'

His greatest single change was from instinct-theory to the analysis of the ego, yet he could not at that stage realize all that this involved, and to the end he tried to ride both horses at once. This is one of the greatest sources of confusion in psychoanalytic theory. But many other important changes occurred, such as the shift of interest from hysteria to obsessional neurosis, and from sex to aggression, the changing concepts of the ego, of anxiety, and of guilt, and many more. Dr Fine seeks to pin-point and expound all this in detail. He shows that Freud has been criticized both for views he did not hold (as that sex is the only source of psychic energy) and for earlier views he had abandoned (as that neuroses are due solely to lack of sexual satisfaction). He shows also that Freud left many concepts of great importance inadequately defined and explored; in particular 'transference' and 'resistance', even though he said that 'any investigation resting on these two facts can call itself psychoanalysis, though it arrives at results other than my own'. This reminds us that what is of far more importance than Freud's actual theories (which, like all scientific theories, belonged to a certain age of thought and must become outdated) is that he provided the starting-points for all genuine psychodynamic inquiry and the observations that enabled his successors to go beyond his own work, and also a seed-bed of ideas that will remain of permanent value. Dr Fine provides excellent summaries of his main works which can be used either as introductions or as memory-refreshers, in both cases leading to the reading or re-reading of Freud himself. H. GUNTRIP

Psycho-Analysis and Faith: the Letters of Sigmund Freud and Oskar Pfister. Ed. by H. MENG and E. L. FREUD. Translated by E. MOSBACHER. London: Hogarth. 1963. Pp. 152. 21*s*.

This interesting book is misnamed. It is a collection of letters of Freud to the Swiss pastor Pfister, with a few from Pfister to Freud, from 1909 to 1938. The question of psychoanalysis and faith is practically confined to pages 110–133, and concerns their discussion of Freud's *The Future of an Illusion* and more briefly *Civilization and its Discontents*. This material is necessarily sketchy and incomplete, for such an enormous subject. It is deeply interesting and contains much of importance on both sides of the argument. But it cannot be more than incidental. Freud says that *The Future of an Illusion* is not part of psychoanalysis but an expression of his private views. The chief interest of this book must be personal and human, not theoretical and philosophical.

Here are two men quite different in temperament and outlook, united in their understanding of psychoanalysis, and in their fundamentally sympathetic natures. Yet they are poles apart in their philosophy of life and their attitudes to 'Man' and 'History'. Freud, the pessimistic and atheist Jew who speaks of mankind as 'trash', reveals delightful casual glimpses of his happy family life, and his deep and even affectionate sympathy for a very difficult young male patient can be seen growing through a series of letters. Pfister, a striking personality, a man of great intelligence, courage and clear views, champions Freud's discoveries in psychological science, feeling great personal respect and friendship for Freud, yet never minimizes for a second the great differences between them once discussion passes beyond psychoanalysis to wider issues. He is thoroughly able to hold his own in the debate, being as well equipped in philosophy as was Freud in science, and to the end regarding Freud as in spirit a better christian than many nominally so-called.

The publication of *The Future of an Illusion* would certainly have caused a rift in the friendship of lesser men. Pfister receives it calmly, replies to it vigorously, discusses Freud's counter-argument, and the two remain as firm friends as before. The great value of this book is to those who wish to understand Freud realistically as a very great human being, as he moved through the early struggles over psychoanalysis, the sinister shadows of the First World War and the rise of Nazism, and his physical sufferings in old age. A pessimist and an agnostic, his antiseptic 'faith' in the importance of work, truth, and so of life itself, never wavered to the very end.

H. GUNTRIP

Critical Essays on Psycho-Analysis. Ed. by S. RACHMAN. Oxford and London: Pergamon. 1963. Pp. xiii + 284. 45*s*.

This book is a collection of ten previously published essays and lectures which have only one thing in common, a critical and in several cases hostile attitude to psychoanalysis. 'Infant Care and Personality' by Orlans, is a useful survey of attempts to test the truth of theories of the effect of 'nursing disciplines' on personality. The conclusions seem to be that it is extremely difficult to devise satisfactory tests, and that it is useless to test any particular discipline (such as breast-feeding, weaning, etc.) by itself. 'The effect of a particular discipline can only be determined from knowledge of the parental attitudes associated with it, the value which the culture places upon that discipline, the organic constitution of the infant, and the entire sociocultural situation in which the individual is located.' It is just this tremendous complexity of the human situation which makes the testing of therapeutic hypotheses so difficult.

Of the other papers, that by A. E. Ellis, 'An Introduction to the Principles of Scientific Psychoanalysis' will best repay study. On many fundamentals a number of the papers contradict each other. Dallenbach's 'Phrenology versus Psychoanalysis' is simply a waste of time. Of the two accounts of personal analyses, Landis took analysis for intellectual, not emotional reasons, and this just does not work; while Boring expected a miracle in 168 sessions and naturally it did not occur.

Percival Bailey spent thirty years as a neurosurgeon, followed by three years 'reading, listening and observing' in psychiatry. He then became 'completely disillusioned' about the 'great psychiatric revolution' he was told had occurred. He enjoys laying about him and slaughtering in succession neurosurgery, E.C.T., drugs and psychoanalysis, and concludes that no real advance has occurred anywhere. However, the scientific mentality is not inconoclastic but researches patiently in face of difficulty. He reaches two contradictory conclusions. The first is that 'the demonstration of a chemical factor in schizophrenia would not help us to understand the contents of schizophrenic delusions: it would merely make them superfluous'. Unfortunately, the attempt to explain mental experience, in general, by chemical factors makes the whole mental life superfluous and we are back in T. H. Huxley's 'epiphenomenon' theory of the mental personality. It is odd then to find Bailey also saying: 'The secret of success of psychotherapy' according to 'a considerable body of convergent opinion' (Ferenczi, Ross, Charcot, Dejerine are cited) lies in 'the function of the personal relationship between the physician and patient'. He omits to note that psychoanalysis has developed into the study of the psychodynamics of just that, therapeutic and human relations.

Unlike Bailey, however, H. J. Eysenck has no use for psychotherapy at all. He produces his familiar 'conditioned reflex' to the stimulus word 'psychoanalysis', the same set-piece with the same examples, arguments and dogmatic opinions. We could all be very happy if there were no more in psychoneurosis than what could be explained by, say, the accidental association of the fear of noises with the sight of white rats. In fact, patients suffer from profound lack of faith in themselves, a sense of the meaninglessness of life and the futility of existence, because they grew up in families where they were not genuinely wanted and had an existence that was barely worth having. Eysenck puts forward Behaviour Therapy with dogmatic assurance as a total replacement for psychodynamic understanding, but he ignores Dr Wynne Jones's statement that it has not yet been developed to the point even where it could be routinely applied, and that its successes in suppressing symptoms often do not relate to the trouble for which the patient was wanting help.

In what sense can psychodynamic studies be held to be scientific? Certainly not in Eysenck's concept of science, but this is the basic theoretical question. Freud is vulnerable because, living when he did, and having the task of establishing psychodynamics as a field of investigation, he was himself inevitably confused as to whether it was science and if so in what sense. What these critics attack is not the psychoanalysis of today, as seen, for example, in the work of Winnicott, or Bowlby or Balint, but the psychoanalytic orthodoxy of thirty years ago. But if every proposition of Freud were to be invalidated with the passage of time, that would make no difference to the fact that he opened up this field of enquiry, and gave the first fruitful leads in the study of it. Its existence does not depend on the truth of Freud's own pioneer theories about it, even though there is, no doubt, a core of truth in his theories which will live. It is the kind of knowledge on which psychotherapy must be based, and it is different from the kind of scientific theory that interests Eysenck, as Dr E. H. Hutten shows very clearly. This finely produced book gives us plenty of negative criticism but little that is constructive towards the solution of psychodynamic problems.

H. GUNTRIP

Handbook of Counselling Techniques. Ed. by ERNEST HARMS and PAUL SCHREIBER. Oxford and London: Pergamon. 1963. Pp. vii + 506. 84*s*.

It was a good idea to attempt a handbook on counselling techniques, for such a book is badly needed. The field is so vast, however, that a single volume has either to be selective in its coverage, of sketchily comprehensive. This book is the latter, having attempted to include all counselling situations except Investment, Consumer and Travel Counselling.

The editors have apparently allowed their contributors considerable freedom in the length of their contributions, and in the style of writing. Thus the long essay on 'The Rabbi as Counsellor' has no subheadings whatever, while other chapters are chopped up by headings and subheadings into short paragraphs like a text-book for junior nurses. Some contributors have given long and authoritative accounts of their own type of counselling, others have done little more than state that they engage in giving other people advice.

This is surely the sort of book in which one would expect some theoretical principles of counselling to be advanced, and possibly some discussion of the variety of roles that are played in counselling. The two pages devoted to general principles in this book, although sound as far as they go, are woefully inadequate. Although principles of counselling are discussed here and there by various authors with reference to their own special type of work, an adequate general discussion at the beginning of the book would have been valuable.

Despite these criticisms, however, this book is important and timely, and will provide a valuable source of data for anyone engaged in counselling work.

RALPH HETHERINGTON

Being-In-The-World. By LUDWIG BINSWANGER. Translated and with a critical introduction to his 'Existential Psychoanalysis' by JACOB NEEDLEMAN. New York and London: Basic Books. 1963. Pp. ix + 364. 52*s.* 6*d.*

Sanity, Madness and the Family. Volume I. *Families of Schizophrenics.* By R. D. LAING and A. ESTERSON. London: Tavistock. 1964. Pp. viii + 272. 30*s.*

Being-in-the-World represents a significant contribution to the literature of psychiatry and of philosophy as it is the first translation into English of Binswanger's writings. The translation by Prof. Needleman affords the reader some weighty selections from Binswanger and these readily whet the appetite for more, e.g. for an English version of his treatise on love, 'Grundformen und Erkenntnis menschlichen Daseins'. Almost half of the present volume is devoted to a lengthy introduction by the translator; this serves to place Binswanger's work within the general perspective of philosophic thought, and some comparison is made with Freud's views. Needleman emphasizes that Binswanger is complementary rather than antagonistic to Freud, and one of Binswanger's essays ('Freud's Conception of Man in the Light of Anthropology') confirms the friendship which he enjoyed with the father of psychoanalysis. Certainly Binswanger and Freud differed over their respective emphasis upon 'spirit' and 'instinct', but the two giants did not turn their backs upon each other as Freud and Jung did ultimately.

Binswanger espouses a 'Verstehen' (understanding) psychology in contradistinction to the 'Erklarung' (explaining) psychology of Freud. He rather deplores Freud's concept of man as 'homo natura' in whom various qualities and values, e.g. goodness, are regarded as derivatives of his instinctual life. Binswanger prefers to regard goodness as an expression of the Dasein, the spiritual sense of 'being-in-the-world'. It would appear that psychotherapy based upon either Freud's or Binswanger's concepts has the common aim of helping the patient face the truth in himself and to confront the inexorability of death. The *a priori* concept of the Dasein owes much to Heidegger's work. The Dasein (being-in-the-world) constitutes a philosophic postulate which is claimed to be necessary in humanising the biological psychiatry of Freud. The Dasein concept underlies the temporality of human existence.

There is an interesting section on schizophrenia which is regarded daseinsanalytically by Binswanger as a primary disorder of the 'self' rather than in terms of Bleulerian symptomatology. The basic problem in schizophrenia is hypothesized as a resignation on the part of the Dasein from its decisional frame of reference, with a consequent splitting of experience into an 'over-reductive' either/or modality. Schizophrenia is thus viewed daseinsanalytically as an impoverished form of existence rather than as a disease. Reality is dealt with by the schizophrenic in an excessively one-sided or 'unproblematic' manner. This philosophic version of schizophrenia is consistent with Laing and Esterson's views. It is a view, however, which is likely to bear less fruit in the field of further research than in that of therapy.

Binswanger emerges from this rather wordy book as a warm, humanistic figure whose existential viewpoint adds a fresh dimension to psychoanalytic theory. At a later date it would be enlightening to learn his view on the affective disorders. Meantime we are in debt to Prof. Needleman for bringing to our attention the work of this eminent authority.

Laing and Esterson have produced a provocative work which has already excited considerable debate and comment, mainly in the pages of *New Society*. Starting from a basic psychoanalytic viewpoint, they have ranged far beyond the conventional limits of analytic theory in this study of family interactions in a selected group of schizophrenic patients. They have carried out careful investigations of the complex and subtle intra-familial relationships which were considered to be relevant to the schizophrenic illnesses of the patients studied.

A detailed account is given of eleven hospitalized female schizophrenic patients aged 15–40 years and of their family relationships. Interviews were carried out, not only with the patients themselves, but also with members of their families in varying combinations and permutations. Much of the material was tape-recorded and subsequently analysed. The diagnosis of schizophrenia in each case was not made by the writers, but had been previously reached by at least two senior psychiatrists. Any patients suspected of having an associated organic cerebral disorder were excluded from the investigation. The method entailed a phenomenological descriptive approach and in this respect the concepts of Jean-Paul Sartre are frequently utilized by the authors in interpreting the social-interactional processes in the patients' families. In the sense of considering the patients' illnesses as being intelligible in the light of environmental processes, the work represents a contribution to social psychiatry.

The case data are vivid and well-written, if on occasions bluntly colloquial in tone. Arguments are adduced to the effect that well recognized schizophrenic symptoms—thought disorder, paranoid delusions, hallucinations, etc.—are understandable as the outcome of highly ambiguous, contradictory and mystifying family situations, especially those involving an unsatisfactory relationship between the patient and her exacting rigid or dominating parents who are reluctant to grant her normal status and automomy. Indeed the illness itself is regarded as a strategic method, devised by the patient, of achieving a degree of self-emancipation or autonomy. Schizophrenia is envisaged not as a disease but as a device for living in a hopelessly confusing environment.

By their method the authors claim to make sense of what in their opinion clinical psychiatrists regard as nonsense. The claim is in many ways attractive since it attempts to convert the bizarre into the intelligible. It should be understood, however, that this point of view is not entirely new: parallels exist in the work of Adolf Meyer, Harry Stack Sullivan, J. N. Rosen and, as indicated above, that of Binswanger and E. Minowski. There is a certain self-satisfaction implicit in the formulations of Drs Laing and Esterson which constitutes a short-coming in their work. An unwarrantable assumption appears to have been made that a one-to-one relationship exists between the occurrence of a specific type of family constellation and the process of schizophrenia. If the authors' assertion is wholly correct, we should require to know why the patients develop schizophrenia rather than a psychoneurosis, a psychosomatic disorder, a personality disorder, delinquency or perhaps no recognizable clinical or social abnormality at all. It would appear rather naïve to maintain that because a disorder may be understood or explicable at the level of disturbed personal communications, we can therefore summarily dismiss genetic, biochemical or other explanations which previously had a scientific status. Surely it is premature to found a new theory, while banishing previous theories, of schizophrenia, on the basis of eleven, admittedly well investigated, female cases. The next step logically required in the authors' scientific journey would be to examine suitable control groups in precisely the same way in which the schizophrenic families have been studied.

A. B. SCLARE

Ursprung und Entfaltung der Seele. By E. VON EICKSTEDT. Stuttgart: Enke. 1963. Pp. viii + 1131. Price DM. 218.

This book, which is subtitled 'A plan and system of a psychological anthropology', is the philosophical and psychological volume of a large three-volume work on anthropology. The author and the publishers feel that this volume is of sufficient interest to be published separately. Prof. Eickstedt is an anthropologist, who rejects both the narrow German philosophical anthropology and the specialized Anglo-American cultural anthropology. For him 'Anthropos is the living human being and anthropology is, therefore, the science of the whole human living being', so that anthropology includes human anatomy, physiology, psychology and sociology.

The theme of the book is the origin and development of man as a psychosocial being. Like the typical painstaking German savant Eickstedt begins with first principles, in this case, atomic physics. He then passes on to neurophysiology, then to psychology and finishes with a discussion of the social and philosophical aspects of his subject. He is mainly dependent on German sources, so that the psychological sections of the book owe a great deal to *Ganzheits*, *Gestalt* and the more philosophical varieties of psychology. Nevertheless, these sections are stimulating and thought provoking.

The amount of literature surveyed by the author is enormous. Some idea of the range of this book can be obtained from the fact that there are 1257 footnotes, each of which cites one or more article or book. Many of the sections of it, although clearly connected to the main theme, are able to stand on their own as excellent short reviews of certain aspects of psychology. For example, the section on 'The ego and the delusional syndrome' is a very good summary of the psychology of the self, which is often neglected by English speaking psychologists. This book is, in fact, unique and every psychologist with a knowledge of German will find it interesting and stimulating. It is very well produced and contains a large number of fascinating and excellent illustrations.

FRANK FISH

A Current Appraisal of the Behavioral Sciences. By Rollo Handy and Paul Kurtz. Great Barrington, Mass.: Behavioral Research Council. 1964. Pp. 154. $7.50.

This survey is published as a Bulletin of its sponsoring Council, whose concern is 'scientific inquiry into the problems of men in Society'. It is itself an interdisciplinary inquiry, the need for which is stressed in some introductory 'General Comments on behavioural research'. Thereafter the work is divided into two sections: 'The Older Fields' and 'The Newer Fields'. The former comprise Anthropology, Sociology, History, Economics, Political Science, Jurisprudence, Psychology and Education; the latter, Information Theory, Cybernetics, Linguistics, Sign-Behaviour (grouped together under the general heading *Communication Theory*), Game Theory, Decision-Making Theory, Value Inquiry (classed under *Preferential Behaviour*) and General Systems.

Each chapter is uniformly structured under nine headings: (1) Working specification of the field. (2) Other specifications of the field. (3) Schools, Methods, Techniques (in Psychology the main stress is on Schools). (4) Results achieved. (5) Contemporary controversy (for Psychology this is perhaps the least satisfactory section). (6) Problems of terminology. (7) Comment and evaluation. (8) Selected Bibliography. (9) Germane Journals (in the case of Psychology all American, with the exception of this *Journal*—although under Information Theory a *British Journal of Psychology and Statistics* is mentioned—and the *Canadian Journal of Psychology*).

To attempt to appraise the appraisal of fields other than one's own is of course impossible; indeed the treatment in all cases is so condensed—while remaining lucid—that it would be idle to try to review even one's 'own' chapter. For each chapter help was given by a number of 'leading scientists': for Psychology they were Hadley Cantril, Alphonse Chapanis, Henry V. Cobb and Robert D. Meade. Nevertheless, the authors are careful to note that the final responsibility rests with themselves and not with the consultants.

How long this *Appraisal* will remain 'current' remains to be seen. As judged by the Bibliography it seems already to be a little dated: the most recent of the 26 books in Psychology was published in 1961, and only two others in 1960. It will be of particular interest to see, should another survey be published, say in ten years' time, whether any of the 'Newer Fields' have established themselves as firmly as the older studies. Psychologists will find satisfaction in the amount of cross-fertilization between their own discipline and several of these Newer Fields, the 'Germane Journals' for which include many belonging primarily to Psychology. The *Appraisal* furnishes hopeful pointers towards the integration of the behavioural sciences, and it looks as if Psychology could play a major part in this process.

B. SEMEONOFF

ERRATUM

Vol. 55, part 1, p. 116.

The authors of *Appraising Vocational Fitness by means of Psychological Tests* are D. E. Super and J. O. Crites.

Brit. J. Psychol. (1964), 55, 4, *pp.* 507–510
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review.)

Adaptive Control Processes: a Guided Tour. By Richard Bellman. Princeton: Princeton University Press; London: Oxford University Press. 1961. Pp. xvi + 255. 42*s*.

The Psychology of Aggression. By Arnold H. Buss. London and New York: John Wiley. 1961. Pp. x + 307. 60*s*.

Borderline Problems in Psychology and Physiology. By E. I. Boiko. Moscow: Academy of Pedagogical Sciences. R.S.F.S.R. 1961. Pp. 212. 83 kop.

Recent Soviet Psychology. Edited By N. O'Connor. Oxford: Pergamon Press. 1961. Pp. 334. 70*s*.

Problema Psicologico dello Spazio. By Enzo Bonaventura. Firenze: Felice Le Monnier. Pubblicazioni della Universita degli Studi de Firenze Facolta di Magistero. 1961. Pp. xii + 146. 2000L.

Training Research and Education. Edited by Robert Glaser. Pittsburgh: University of Pittsburgh Press. 1962. Pp. xii + 596. $11.00.

Handbook of Small Group Research. By A. Paul Hare. New York: The Free Press of Glencoe. 1962. Pp. xiv + 512. $10.00.

Individual in Society. By David Krech, Richard S. Crutchfield and Egerton L. Ballachey. London: McGraw-Hill. 1962. Pp. xii + 564. 62*s*.

Industrial Psychology. By Laurence Siegel. Homewood, Illinois: Richard D. Irwin. 1962. Pp. xvi + 414. $6.95.

Frustration and Conflict. By Aubrey J. Yates. London: Methuen. 1962. Pp. x + 236. 28*s*.

Transfer and Communality. By Karl-Georg Ahlström. Uppsala: Almqvist and Wiksells. 1961. Pp. 156. N.P.G.

Erinnerungen, Träume, Gedanken von C. G. Jung. By Aniela Jaffé. Zurich and Stuttgart: Rascher. 1962. Pp. vi. + 422. sFr. 29.

L'Année Psychologique (1962), nos. 1 and 2. Paris: Presses Universitaires de France. 1962. Pp. xii + 342; Pp. viii + 702. F. 25 + T.L.; F. 28 + T.L.

L'Année Psychologique (1963), nos. 1 and 2. Paris: Presses Universitaires de France. 1963. Pp. vii + 278. F. 25 + T.L.; F. 28 + T.L.

The Clinical Use of Dreams. By Walter Bonime. London: Basic Books. 1963. Pp. xxvii + 343. $8.50.

Conflicts of the Clergy. By Margaretta K. Bowers. New York: Nelson. 1963. Pp. xvi + 252. $4.95.

Educational Psychology (2nd edition). By Lee J. Cronbach. London: Hart-Davis. 1963. Pp. xxvii + 706. 65*s*.

The Diagnosis and Treatment of Speech and Reading Problems. By Carl H. Delacato. Illinois: Thomas. 1963. Pp. x + 188. $6.75.

Man's Image in Medicine and Anthropology. Edited by Iago Galdston. New York: International Universities Press. 1963. London: Bailey Bros. and Swinfen. 1963. Pp. xvii + 525. 90*s*.

The Paradox of Pleasure and Relativity: the Psychological Causal Law. By D. G. Garon. New York: Philosophical Library. 1963. Pp. x + 499. $6.00.

This is the Law: The Magistrates' Court. By F. T. Giles. London: Stevens. 1963. Pp. viii + 250. 22*s*. 6*d*.

A Psychiatrist's Views on Social Issues. By Sol W. Ginsburg. London: Columbia University Press. Pp. xviii + 296. 37*s*. 6*d*.

An Essay on Eyesight. By Oliver Jelly. London: Hodder and Stoughton. 1963. Pp. 160. 35*s*.

The Psychology of Adolescence (*2nd edition*). By Arthur T. Jersild. New York: Macmillan; London: Collier-Macmillan. 1963. Pp. xi + 468. N.P.G.

Personal Adjustment: an Approach through the Study of Healthy Personality (2nd edition). New York: Macmillan; London: Collier-Macmillan. 1963. Pp. xiii + 477. N.P.G.

Mysterium Oniunctionis. By C. G. Jung. Vol. 14—*The Collected Works.* London: Routledge and Kegan Paul. 1963. Pp. xix + 704. 52*s*. 6*d*.

Personal Responsibility: Growth and Limits. By Monica Lawlor. London: Burns and Oates. 1963. Pp. 142. 9s. 6*d*.

Psychiatric Disorder Among the Yoruba. By A. H. Leighton, T. A. Lambo, C. C. Hughes, D. C. Leighton, J. M. Murphy and D. B. Macklin. London: Oxford University Press. 1963. Pp. xix + 413. 70*s*.

Lysergic Acid (*LSD* 25) *and Ritalin in the Treatment of Neurosis.* By T. M. Ling and J. Buckman. Sidcup, Kent: Lambarde Press. 1963. Pp. 172. 21*s*.

Syndromes of Psychosis. By M. Lorr, C. J. Klett and D. M. McNair. Oxford: Pergamon Press. 1963. Pp. x + 286. 70*s*.

The Mentally Retarded Child. Edited by A. R. Luria. Oxford: Pergamon Press. 1963. Pp. viii + 207. 50*s*.

Child Development and Personality (2nd edition). By P. H. Mussen, J. J. Conger and J. Kagan. New York: Harper and Row. 1963. Pp. xiv + 625. 55*s*.

Alfred Adler: The Man and His Work. Triumph over the Inferiority Complex. By Hertha Orgler. London: Sidgwick and Jackson. 1963. Pp. xix + 266. 30*s*.

Soundmaking: the Acoustic Communication of Emotion. By Peter F. Oswald. Springfield, Illinois: Charles C. Thomas. 1963. Pp. xii + 186. $7.50.

Metamorphosis. On the Development of Affect, Perception, Attention, Memory. By Ernest G. Shachtel. London: Routledge and Kegan Paul. 1963. Pp. viii + 344. 35*s*.

Narcotic Addiction in Britain and America. By Edwin M. Schur. London: Tavistock Publications. 1963. Pp. 281. 35*s*.

Pediatric Psychiatry. By Hale F. Shirley. Harvard: University Press; London: Oxford University Press. 1963. Pp. xx + 796. 80*s*.

Theory of Relationships. By Sanford L. Silverman and Martin G. Silverman. New York; Philosophical Library. 1963. Pp. xv + 111. $6.00.

Mind and Consciousness in Experimental Psychology. By R. H. Thouless. Cambridge University Press. 1963. Pp. 32. 3*s*. 6*d*. (The seventeenth Arthur Stanley Eddington Memorial Lecture, 29 November 1963).

Frontiers of Psychology. By G. K. Yacorzynski. New York: Philosophical Library. 1963. Pp. x + 229. $5.00.

Autobiographies of Three Pagans in the Philippines. By R. F. Barton. New York: University Book, Inc. 1963. Pp. xliii + 267. $7.00.

Sittings with Eusapia Palladino and Other Studies. By Everard Feilding. New York: University Books, Inc. 1963. Pp. xx + 324. $10.00.

From India to the Planet Mars. By Theodore Flournay. New York: University Books, Inc. 1963. Pp. xxxvi + 457. $10.00.

Apparitions and Precognition. By Aniela Jaffé. New York: University Books, Inc. 1963. Pp. viii + 214. $7.50.

Development of Psychological Thought in India. By S. K. Ramachandra Rao. Mysore: Kayvalaya Publishers. 1963. Pp. ix + 225. 28*s*.

Obituary: The 'Hodgson Report' on Madame Blavatsky—1885–1960. By Adlai E. Waterman. Madras: the Theosophical Publishing House. 1963. Pp. xx + 92. 2·50 Rupees.

Abhandlungen zur Religions-und-Arbeits Psychologie. By Alfons Bolley and Gerhard Clostermann. Münster: Aschendorff. 1963. Pp. 88. DM. 7.50.

Zur Psychologie des produktiven Denkens. By Karl Duncker. Berlin: Springer. 1963. Pp. vii + 135. N.P.G.

Das Verbrechen: III Anlage-Komponenten im Getriebe des Delikts. By Hans von Hentig. Berlin: Springer. 1963. Pp. viii + 523. DM. 59.

Die schizophrenieähnlichen Emotionspsychosen. By F. Labhardt. Berlin: Springer. 1963. Pp. vi + 103. DM. 28.

Über den methodischen Fortschritt in der Umfrageforschung. By Elizabeth Noelle. Allensbach, West Germany: Verlag für Demoskopie. 1963. Pp. 69. DM. 9.50.

Psychiatrie im Grundriss. By Hans Jorg Weitbrecht. Berlin: Springer. 1963. Pp. xvi + 486. DM. 36.

La Filiation des Structures. By L. Apostel, J. B. Grize, S. Papert and J. Piaget. Paris: Presses Universitaires de France. 1963. Pp. iv + 196. NF. 9.50 + T.L.

La Formation des Raisonnements Récurrentiels. By P. Greco, B. Inhelder and J. Piaget. Paris: Presses Universitaires de France. 1963. Pp. 321. 14.50 F.

Sociologie et Psychologie. By René Duchac. Paris: Presses Universitaires de France. 1963. Pp. 123. F 5.

Perception Tachistoscopique et Communication. By A. Fauville. Louvain: Editions Nauwelaerts. 1963. Pp. 52. 80 FB.

La Perception. By Robert Francès. Paris: Presses Universitaires de France. 1963. Pp. 128. N.P.G.

Les Gauchers: Prévalence Manuelle et Dominance Cérébrale. Paris: Presses Universitaires de France. 1963. By Henry Hécaen and Julian de Ajuriaguerra. Pp. 171. F. 12 + T.L.

Le Sourire du Nourrisson. La Voix Comme Facteur Declenchant. By J. L. Laroche and F. Tcheng. Louvain: Publications Universitaires. 1963. Pp. 184. N.P.G.

Communications et Langages. Edited by A. A. Moles and Bernard Vallancien. Paris: Gauthier-Villars. 1963. Pp. xv + 215. 40 F.; $8.50.

Modèles Sociométriques et Formation des Cadres. By Roger Mucchielli. Paris: Presses Universitaires de France. 1963. Pp. 156. F. 18 + T.L.

Examens et Docimologie. By Henri Piéron. Paris: Presses Universitaires de France. 1963. Pp. 190. F. 9.

La Conscience. By Henri Ey. Paris; Presses Universitaires de France. 1963. Pp. 439. F 20 + T.L.

Les Activités Intellectuelles. By Pierre Oléron. Paris: Presses Universitaires de France. 1964. Pp. 163. F 9 + T.L.

Typological properties of Higher Nervous Action in Man, vol. iii (in Russian). Edited by B. M. Teplov. Moscow: Academy of Pedagogical Sciences of the R.S.F.S.R. 1963. Pp. 276. 1 r. 19 kop.

History of Russian Psychology. Edited by M. V. Sokolov. Moscow: Academy of Pedagogical Sciences, R.S.F.S.R. 1963. Pp. 439. 1 r. 08 kop.

The Uses of Psychology in Geriatric Ophthalmology. By WERNER BAB. Springfield, Illinois: Thomas. 1964. Pp. x + 94. $5.00.

Advances in Sex Research. Edited by HUGO G. BEIGEL. London: Harper and Row. 1964. Pp. xiv + 261. 52*s*.

Psycho-Therapy: a Christian Approach. By E. N. DUCKER. London: George Allen and Unwin. 1964. Pp. 126. 21*s*.

Mankind Behaving: Human Needs and Material Culture. By JAMES K. FEIBLEMAN. Springfield, Illinois: Charles C. Thomas. 1963. Pp. vii + 275. $10.50.

Counterpoint: Libinal Object and Subject. Edited by HERBERT S. GASKILL. New York: International Universities Press. 1963. Pp. 200. 40*s*.

On the Theory of Social Change. By EVERETT E. HAGEN. London: Tavistock. 1964. Pp. xvii + 557. 63*s*.

Changing Parental Attitudes. By CARL F. HEREFORD. Austin, Texas: University of Texas Press. 1964. Pp. xvi + 198. $5.00.

The Early Growth of Logic in the Child. By BARBEL INHELDER and JEAN PIAGET. London: Routledge and Kegan Paul. 1964. Pp. xxv + 302. 40*s*.

A Two-Year-Old goes to Nursery School. By MARJORIE GRAHAM JANIS. London: Tavistock. 1964. Pp. xi + 156. 25*s*.

The Meaning of History. By ERICH KAHLER. New York: George Braziller. 1964. Pp. viii + 224. $5·00.

A History of the Care and Study of the Mentally Retarded. By LEO KANNER. Springfield; Illinois: Thomas. 1964. Pp. x + 150. $6.50.

Reason, Action and Morality. By J. KEMP. London: Routledge & Kegan Paul. 1964. Pp. viii + 208. 30*s*.

The Universal Experience of Adolescence. By NORMA KIELL. New York: International Universities Press. 1964. London: Bailey Bros. and Swinfen. Pp. 942. 112*s*.

Lectures in Dynamic Psychiatry. Edited by MILTON KURIAN and MORTON H. HAND. New York: International Universities Press. 1964. Pp. 137. 27*s*.

Conditioning and Enuresis. By S. H. LOVIBOND. Oxford: Pergamon Press. 1964. Pp. viii + 219. 50*s*.

Society in the Mind: Elements of Social Eidos. By CHARLES MADGE. London: Faber and Faber. 1964. Pp. 158. 25*s*.

Hypnosis Throughout the World. Edited by F. L. MARCUSE. Springfield, Illinois: Thomas. 1964. Pp. xi + 312. $11.00.

Problems of the Biochemistry of the Nervous System. By A. V. PALLADIN. Oxford: Pergamon Press. 1964. Pp. xii + 330. 80*s*.

Lectures on Conditioned Reflexes. By IVAN P. PAVLOV. London: Lawrence and Wishart. 1964. Vol. I. Pp. 414, 42*s*. Vol. II. Pp. 199, 21*s*.

Working with Children in Hospitals. By EMMA A. PLANK. London: Tavistock. 1964. Pp. v + 86. 21*s*.

The Genain Quadruplets: a Study of Heredity and Environment in Schizophrenia. Edited by DAVID ROSENTHAL. London: Basic Books. 1964. Pp. xiv + 609. 84*s*.

Indian Students in Britain: A Survey of their Adjustment and Attitudes. By AMAR KUMAR SINGH. London: Asia Publishing House. 1964. Pp. xx + 208. 45*s*.

The Divided Self. By W. S. STEWART. London: George Allen and Unwin. 1964. Pp. 246. 40*s*.

An Introduction to Psychotherapy. By SIDNEY TARACHOW. London: Hogarth. 1964. Pp. viii + 376. 35*s*.

Psychiatric Social Work in Great Britain (1939–1962). By Noel Timms. London: Routledge and Kegan Paul. 1964. Pp. ix + 270. 32*s*.

Psychiatry on the College Campus. By H. G. Whittington. New York: International Universities Press; London: Bailey Bros. and Swinfen. 1964. Pp. 328. 68*s*.

Homer Lane. By W. David Wills. London: George Allen and Unwin Ltd. 1964. Pp. 275. 40*s*.

Buddha and the Gospel of Buddhism. By Ananda Coomaraswamy. New York: University Books. 1964. Pp. ix + 369. $10.00.

The Case of Patience Worth. By Walter Franklin Prince. New York: University Books Inc. 1964. Pp. 509. $10.00.

The Mediumship of Mrs Leonard. By Susy Smith. New York: University Books. 1964. Pp. 260. $7.50.

Psychiatric Education: a Symposium held at the Institute of Psychiatry. The Maudsley Hospital, London on 14th–15th March 1963. London: Pitman Medical Publishing Co. 1964. Pp. 129. 15*s*.

Magical Thinking, Modern Medicine and the Curriculum. By Wallace Ironside. Dunedin: University of Otago Press. 1963. Pp. 20. N.P.G.

Perception of Parental Identification, Parental Dominance and Anxiety in Young Adults. By Sven Marke and G. Eberhard Nyman. Lund, Sweden: Ohlsson. 1963. Pp. 47. Kr. 6.

Der Tuberkulöse im Strafvollzug. By Th. Johannes. Berlin: Springer Verlag. 1964. Pp. vii + 196 N.P.G.

L'Organisation Perceptive. By Eliane Vurpillot. Paris: Librarie Philosophique. J. Vrin. 1963. Pp. 188. N.P.G.

Phenomenology of Perception. By M. Merleau-Ponty. London: Routledge and Kegan Paul. 1962. Pp. xxii + 466. 46*s*.

Action, Emotion and Will. By Anthony Kenny. London: Routledge and Kegan Paul. 1963. Pp. viii + 245. 25*s*.

Sensationalism and Scientific Explanation. By Peter Alexander. London: Routledge and Kegan Paul. 1963. Pp. viii + 149. 16*s*.

The Causation of Bus Driver Accidents. By W. L. Cresswell and P. Froggatt. London: Oxford University Press. 1963. Pp. xvii + 298. 45*s*.

The Practice of School Psychology: professional problems. By Robert E. Valett. London and New York: Wiley. 1963. Pp. x + 339. 57*s*.

The Structure and Dynamics of Organisations and Groups. By Eric Berne. London and Montreal: Pitman. 1963. Pp. x + 260. 60*s*.

Working with Groups. By Josephine Klein. London: Hutchinson. 2nd edition. 1963. Pp. 240. 15*s*.

The Science of Behaviour. By John McLeish. London: Barrie and Rockliff. 1963. Pp. vi + 199. 18*s*. 6*d*.

Scientific Principles of Psychology. By Donald J. Lewis. Eaglewood Cliffs, New Jersey and London: Prentice Hall. 1963. Pp. xi + 580. 56*s*.

Cybernetics: Machines with Intelligence. By Neville Moray. London: Burns and Oates. 1963. Pp. 124. 9*s*. 6*d*.

The Minds of Robots. By James T. Culbertson. Urbana: Illinois University Press. 1963. Pp. 466. $10.00.

Experiments with Drugs. By H. J. Eysenck. Oxford: Pergamon. 1963. Pp. xii + 421. £5.

Progress in Brain Research, Vol. 2. *Nerve, Brain and Memory Models*. Ed. by Norbert Wiener and J. P. Schadé. Amsterdam, London and New York: Elsevier. 1963. Pp. viii + 280. 80*s*.

The Social Adjustment of Children. Manual to the Bristol Social Adjustment Guides. 2nd edition. London: London University Press. 1963. Pp. 63. 25*s.*

Psychiatric Studies of Borstal Lads. (Maudsley Monographs No. 11.) By T. C. N. GIBBENS and others. London: Oxford University Press. 1963. Pp. viii + 230. 45*s.*

The Education of Technologists. An exploratory study at Brunel College. By MARIE JAHODA. London: Tavistock. 1963. Pp. xiii + 226. 30*s.*

Disorders of Language: Ciba Foundation Symposium. By A. V. S. DE REUCK and MAEVE O'CONNOR. London: Churchill. 1964. Pp. xii + 356. 50*s.*

Basic Readings in Neuropsychology. Ed. by ROBERT L. ISAACSON. London: Harper and Row. 1964. Pp. xi + 429. 37*s.*

Healing the Sick Mind. By HARRY GUNTRIP. London: Allen and Unwin. 1964. Pp. 223. 13*s.* 6*d.*

Reason and Violence. By R. D. LAING and D. G. COOPER. London: Tavistock. 1964. Pp. 184. 25*s.*

Biological Basis of Behaviour. A program. By F. J. MCGUIGAN. London: Prentice-Hall. 1964. Pp. xiv + 238. 32*s.*

They Wouldn't Stop Talking. By JOHN POLLOCK. London: Blond. 1964. Pp. 156. 18*s.*